PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C07K 15/28, A61K 37/02
C12N 15/13, C12P 21/08

A1

(11) International Publication Number: **WO 91/16353**

(43) International Publication Date: 31 October 1991 (31.10.91)

(21) International Application Number: PCT/EP91/00767

(22) International Filing Date: 21 April 1991 (21.04.91)

(30) Priority data:
90401090.7 23 April 1990 (23.04.90) EP
(34) Countries for which the regional or international application was filed: GB et al.

(71) Applicant *(for all designated States except US):* CORVAS INTERNATIONAL N.V. [BE/BE]; J. Plateaustraat 22, B-9000 Gent (BE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only) :* LAROCHE, Yves [BE/BE]; 115, rue Bemel, B-1150 Brussels (BE). HOLVOET, Paul [BE/BE]; Broakstraat 13A, B-3010 Kessel-Lo (BE). DE MAEYER, Marc [BE/BE]; Keienveld 2, B-1702 Groot-Bijgaarden (BE).

(74) Agents: GUTMANN, Ernest et al.; Ernest Gutmann-Yves Plasseraud S.A., 67, bd Haussmann, F-75008 Paris (FR).

(81) Designated States: AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH (European patent), CM (OAPI patent), DE (European patent), DK (European patent), ES (European patent), FI, FR (European patent), GA (OAPI patent), GB (European patent), GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU (European patent), MC, MG, ML (OAPI patent), MR (OAPI patent), MW, NL (European patent), NO, PL, RO, SD, SE (European patent), SN (OAPI patent), SU, TD (OAPI patent), TG (OAPI patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: THROMBUS-SPECIFIC ANTIBODY DERIVATIVES

(57) Abstract

The invention relates to fibrin-specific single-chain antibodies, thrombolytic agents derived from such antibodies, and DNA fragments coding for such polypeptides. The single-chain antibodies can be used for imaging, while the thrombolytic agents can be used for *in vivo* lysis of thrombi.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

THROMBUS-SPECIFIC ANTIBODY DERIVATIVES

This invention relates to: novel thrombus (preferably fibrin)-binding molecules ("SCAs"), derived from thrombus (preferably fibrin)-specific antibodies; novel thrombolytic agents ("SCAPAs") derived from SCAs; and genes coding for such SCAs and SCAPAs.

This invention also relates to the uses of the SCAs and SCAPAs and to methods of producing them.

Background of the Invention

An injury to a blood vessel normally results in the activation of the complex hemostatic process (Colman et al., 1987) and the formation of a blood clot at the site of injury. The blood clot is predominantly composed of blood platelets emmeshed in a network of fibrin, and its formation protects the organism from bleeding. To restore normal blood flow, the clot is later remodelled and removed by degradation of fibrin by proteolytic enzymes such as plasmin.

Under certain pathological conditions the hemostatic process can also result in the formation of a thrombus, which is a solid mass or plug formed in a living heart or in blood vessels which can cause severe complications (e.g., myocardial infarctions). A thrombosis is due to either local obstruction of blood vessels or distant embolization. The relative amounts of the formative elements of thrombi (i.e., blood platelets, erythrocytes and fibrin) depend on the place in the vascular system where the thrombi are formed (Freiman, 1987).

Fibrin is formed by polymerization of fibrinogen, a protein with a very complex organization. Each fibrinogen molecule is composed of six complexly linked polypeptides (i.e., 2 A-alpha, 2 B-beta and 2 gamma

polypeptides). The molecule can be divided into a number of domains, the most important of which are the two terminal D-domains and the central E-domain. During fibrin formation, covalent crosslinks are formed between D and D and between D and E domains of adjacent fibrinogen molecules, resulting in a highly interconnected network of fibers with a sufficient mechanical stability to serve as a hemostatic plug (Hantgan et al., 1987). Degradation of fibrin by plasmin results in a variety of degradation products, one of which is the so-called D-dimer which comprises two cross-linked D domains of adjacent fibrinogen molecules.

Therapy of thrombosis and its complications has mainly been directed towards: prevention of local extension of a thrombus by administration of anticoagulants such as heparin; and/or acceleration of the rate of dissolution of the thrombus by administration of thrombolytic agents such as plasminogen activators which are substances that initiate formation of plasmin (a proteolytic enzyme that degrades fibrin). The most important plasminogen activators are tissue-type plasminogen activators ("tPA"), urokinase-type plasminogen activators ("uPA") and streptokinase and its derivatives (Bachmann, 1987).

tPA is a protein of 530 amino acids which can be isolated from many tissues but which seems to be mainly produced by endothelial cells as a single-chain polypeptide. The amino acid sequence of this polypeptide and the nucleotide sequence that encodes it have been described by Pennica et al. (1983). tPA is cleaved by plasmin between Arg278 and Ile279 resulting in a molecule consisting of two polypeptide chains held together by one disulfide bridge. In the presence of fibrin both forms are equally active. Structurally, tPA

can be divided in a finger domain containing low-affinity fibrin binding sites, a domain with homology to epidermal growth factor ("EPG"), two kringle domains with high-affinity fibrin binding sites, and the serine protease catalytic domain. Because of its binding capacity to fibrin, plasminogen activation by tPA is highly fibrin-specific.

An urokinase precursor, designated as prourokinase or single-chain uPA ("scuPA"), is synthesized in tissues such as endothelium of the kidney as a continuous polypeptide of 411 amino acids and a molecular weight of about 54 kiloDalton (kD) ("scuPA-54k"). The amino acid sequence of this polypeptide and the nucleotide sequence that encodes it have been described by Holmes et al. (1985). ScuPA is converted to the active form, designated as urokinase or two-chain uPA ("tcuPA") by a proteolytic cleavage (e.g., by plasmin) between Lysine-158 and Isoleucine-159 which results in a molecule consisting of two polypeptide chains held together by a single disulfide bond. Three domains can be identified in prourokinase: a EPG domain, one kringle domain and a serine protease catalytic domain. These domains show important homologies with the corresponding domains of tPA. However, neither the EPG domain, nor the kringle domain of scupa have affinity for fibrin.

A low molecular weight form of scuPA ("LMW-scuPA") is formed by an additional proteolytic cleavage between lysine 135 and Lys136, releasing the first 135 amino acids (i.e., the EPG and the kringle domains). The LMW-scuPA is converted to a low molecular weight two-chain urokinase of about 33 kD ("LMW-tcuPA") by a proteolytic cleavage between Lys158 and Ile159.

European patent publication ("EP") 247674 describes a single-chain uPA of about 32 kD which

apparently is derived from the scuPA by a proteolytic cleavage between Glu143 and Leu144 and which can be recovered in stable conditions and in fair yields from the culture fluid of human lung adenocarcinoma cells of the type CALU-33 (ATCC cell line HTB-55). This form was designated as "scuPA-32k" to differentiate it from the previously known scuPA-54k. ScuPA-32k, like scuPA-54k, can be proteolytically activated, e.g., by plasmin.

Although scuPA-54k does not bind directly to fibrin, it activates fibrin-associated plasminogen much more readily than plasminogen in the plasma (Gurewich et al., 1984). Although scuPA-32k lacks both the EPG and the kringle domain of scuPA-54k, it still displays fibrin selectivity. In contrast, the two active forms, described above, have lost their fibrin selectivity.

The use of all plasminogen activators as thrombolytic agents has been hampered by the fact that they will also dissolve existing hemostatic plugs in an organism, thus leading to hemorrhaging. Plasminogen activators that lack fibrin affinity, such as streptokinase and urokinase, will also activate circulating plasminogen resulting in impairment of platelet function and degradation of circulating fibrinogen and clotting factors V and VIII (in addition to the fibrin in the thrombus) and leading to systemic effects which have generally been designated as the "lytic state" (Marder and Bell, 1987). In general, it is believed that such systemic effects of plasminogen activator therapy can be minimized by the use of plasminogen activators, such as tPA or scuPA, which demonstrate fibrin-selective plasminogen activation, thus releasing plasmin only in the vicinity of a clot.

tPA and the active forms of uPA have very short circulatory half-lives (5 and 7 minutes for tPA and urokinase, respectively - Sherry, 1987) due to their

interaction with receptors on cells of the liver. In addition, their activatory half-life is reduced by means of a rapid and irreversible inactivation by plasminogen activator inhibitor I (Haber et al., 1989). Both scuPA-54k and scuPA-32k are resistant to this inhibitor and have longer activatory half-lives. Nevertheless, treatment with any of these agents has had to be over extended periods and in combination with heparin to avoid the occurrence of reocclusion.

In attempts to improve on the naturally occurring plasminogen activators, several approaches have been followed (see, e.g., Haber et al., 1989; Haber, 1990). Generally, plasminogen activators have been sought which display, in addition to high thrombus (e.g., fibrin)-selective activity, longer circulatory and activatory half-lives.

Antibodies are proteins that are secreted by specialized cells (i.e., the B-lymphocytes) as a part of the immune response of an organism to introduction of foreign molecules (i.e., antigens). Any antibody is highly specific for one particular antigen. With the development of hybridoma technology, monoclonal antibodies, which are highly homogenous with respect to their antigen specificity, could be produced in virtually unlimited quantities (see, e.g., Harlow and Lane, 1988).

The structure of antibodies is well known (see, e.g., Albers et al., 1989). They have a Y-shape and consist of two identical heavy (H) chains of about 440 amino acids and two identical light (L) chains of about 220 amino acids. The various polypeptide chains of a single antibody molecule are connected to each other by four disulfide bridges. In mammals, there are five different classes of antibodies, each of which is characterized by a different type of H-chain (the

alpha, delta, epsilon, gamma and mu H-chains). In addition, antibodies can contain two different L chains (the kappa and lambda L-chains). The type of H-chain defines effector functions other than antigen specificity, such as interactions with antibody receptor molecules on different cells. The difference between the two types of L-chains has not yet been identified. The largest class of circulating antibodies (IgG) have a gamma H-chain.

The H- and L-chains consist of variable and constant domains. The L-chain has a variable domain of about 110 amino acids ("V_L") and a constant domain also of about 110 amino acids ("C_L"). The H-chain has one variable domain ("V_H") and three constant domains ("C_{H1}", "C_{H2}" and "C_{H3}"), each of about 110 amino acids. In each antibody molecule, there are two antigen binding sites, each of which is formed by the variable domains of one L-chain and one H-chain, more specifically by the so-called complementarity determining regions ("CDRs") within the variable domains, and there are three CDRs on each chain. The amino acid sequences of the CDRs are highly variable among antibodies while the sequences of the parts of the variable domains next to and in between the CDRs (i.e., the so-called "framework regions") are much more conserved. The antigen binding site interacts with a well-defined region of the antigen which is designated the "epitope".

The variable domains of the heavy and light chains are encoded by different gene segments which are properly organized in the fully differentiated B-lymphocyte through recombination events. The V_L region thus consists of a large N-terminal part which is encoded by the so-called variable ("V") gene segment and a short C-terminal part which is encoded by the

so-called joining ("J") gene segment. The V_H domain largely consists of a large N-terminal part, which is encoded by a V gene segment, and two smaller parts, encoded by the diversity ("D") gene segment and J gene segment, respectively.

In principle, the V_L and V_H are the only components necessary for antigen binding. It has been shown that proteins can be prepared by connecting the nucleotide sequences coding for the V_L and V_H regions with a linker sequence coding for a linker polypeptide ("L"), and expression of these hybrid DNA molecules can be obtained in E. coli (Bird et al., 1988; Huston et al., 1988; Chaudhary et al., 1989, 1990). The resulting V_L-L-V_H or V_H-L-V_L proteins retained their antigen-binding capacity and can be designated as single-chain antibodies.

Although general rules have been proposed for designing suitable linkers (see, e.g., PCT publication WO 88/01649), the actual design of a single-chain antibody that retains the affinity and specificity of the original antibody, from which it was derived, and that can be produced in appropriate host cells is far from straightforward.

To increase the fibrin specificity of plasminogen activators, it has been attempted to link the catalytic portions of plasminogen activators with the antigen binding sites of fibrin-specific monoclonal antibodies, for instance by construction of chimaeric molecules with both biological functions. Anti-fibrin antibody 59D8, directed against the amino-terminal six-amino acid sequence of the fibrin beta chain, was chemically conjugated to urokinase and tPA (Bode et al., 1987), and anti-fibrin antibody MA-15C5, directed against human fibrin D-dimer, was conjugated to scupa (Collen et al., 1989; Dewerchin et al., 1990; Collen et al.,

1990). Recombinant DNA technology has also been used to replace parts of the heavy chain of the 59D8 antibody with portions of tPA and uPA catalytic domains (EP 271227, EP 355068 and EP 347078). In general, the results of these attempts have been mixed, and an ideal thrombolytic agent has not yet been identified (Haber et al., 1989; Haber, 1990).

Summary of the Invention

This invention provides an SCA comprising a single-chain antibody which can bind in a highly specific manner to at least one thrombus constituent, preferably fibrin. It is preferred that the SCA comprise all or preferably the effective antigenic-binding parts of a monoclonal antibody directed against the thrombus constituent, particularly fibrin, quite particularly fibrin cross-links, in a thrombus. It is particularly preferred that the SCA comprise all or preferably the effective antigenic-binding parts of the heavy and light variable domains (V_H and V_L, respectively) of the monoclonal antibody, linked through a first linker peptide ("L_{ab}") so as to form a single chain.

This invention also provides a method of using the SCA for imaging of thrombi and for making novel thrombolytic agents comprising the SCA as a thrombus constituent-binding portion, preferably a fibrin-binding portion.

This invention further provides an SCAPA which is a plasminogen activator comprising an SCA as a thrombus-binding portion ("SCA-portion") connected to a plasminogen activating portion ("PA-portion"). It is preferred that the PA-portion comprise at least the catalytic domains of a plasminogen activator, preferably of tPA or uPA, particularly of scuPA. The

C-terminal end of the SCA-portion is preferably directly linked to the N-terminal end of the PA-portion, but both portions can also be linked through a second linker peptide ("L_{cd}").

This invention still further provides a DNA molecule coding for the SCA ("sca gene") or for the SCAPA ("scapa gene"), a chimaeric DNA sequence ("chimaeric gene") containing the sca or scapa gene and a vector containing the chimaeric gene. Preferably, the chimaeric gene comprises the following operably linked DNA fragments in the same transcriptional unit: 1) a promoter capable of directing expression of a sca or scapa gene in a procaryotic or eucaryotic host cell; 2) a sca or scapa gene; and 3) suitable 3' regulatory sequences. The chimaeric gene can optionally contain, between DNA fragments 1) and 2), a signal sequence that encodes a polypeptide ("signal peptide") directing the secretion of the SCA or SCAPA from the procaryotic or eucaryotic host.

This invention further provides a method for obtaining the SCA or the SCAPA by: introducing the chimaeric gene in a procaryotic or eucaryotic host cell so that it is actively expressed within the host cell; culturing the host cell; and then recovering the SCA or SCAPA from the culture.

Brief Description of the Drawings

Fig. 1 - Amino acid sequence of the variable region of the kappa chain of the monoclonal antibody MA-15C5. The numbering of the amino acids follows the generalized numbering described by Kabat et al. (1987). Single lines indicate the borders of the CDR and framework regions.

Fig. 2 - Amino acid sequence of the variable region of the gamma-chain of the monoclonal antibody

MA-15C5. The numbering of the amino acids follows the generalized numbering described by Kabat et al. (1987). Single lines indicate the borders of the CDR and framework regions. The double line indicates the end of the region encoded by the J gene segment.

Fig. 3 - Nucleotide sequence of the cDNA coding for the variable and constant region of the kappa chain of the monoclonal antibody MA-15C5. The amino acid sequence of the variable region is also given (see also Fig. 1). The sequence also comprises the signal sequence. Important restriction sites used during cloning procedures are indicated.

Fig. 4 - Nucleotide sequence of the cDNA coding for the variable region of the gamma chain of the monoclonal antibody MA-15C5 with the amino acid sequence as given in Fig. 2. The codons for the first four amino acids of the V_L domain are missing. The sequence also comprises part of the coding sequence of the C_{H1} domain of the MA-15C5 gamma chain. Important restriction sites used during cloning procedures are indicated. The double line indicates the end of the region encoded by the J gene segment.

Fig. 5 - Nucleotide sequence and deduced amino acid sequence of the human uPA cDNA. Important restriction sites used during cloning procedures are indicated. The cleavage site of LMW-tcuPA (single vertical line) and the N-terminus of scuPA-32k (double vertical line) are indicated.

Fig. 6 - Amino acid sequences of preferred SCAs of this invention. The numbers and amino acids in square brackets refer to the amino acid sequence of V_L of MA-15C5 as given in Fig. 1. The numbers between brackets refer to the amino acid sequence of the first linker peptide (L_{ab}). The letters and amino acids between accolades refer to the amino acid sequence of

V_H of MA-15C5 as given in Fig. 2. The sequences derived from the V_L and V_H anchor regions are underlined. Residues marked with an asterisk are residues that are mutated with respect to the original sequence.

Fig.7 - Nucleotide sequence of the tac promoter and the PhoA signal sequence. The encoded amino acid sequence of the PhoA signal peptide is also shown. The promoter and signal peptide can for instance be used for the expression and secretion of foreign proteins in E. coli.

Fig. 8 - A. Nucleotide sequence of the signal sequence of the kappa chain of MA-15C5 monoclonal antibody. The encoded amino acid sequence of the kappa-chain signal peptide is also shown.

- B. Nucleotide sequence of a signal sequence coding for a consensus signal peptide of a human IgG gamma chain. The 11 C-terminal amino acids of this signal peptide are those of the natural signal peptide of the gamma chain of MA-15C5 monoclonal antibody. The signal peptides of Figs. 8a and 8b can, for instance, be used for the expression and secretion of foreign proteins in eucaryotic cells.

Fig. 9 - Nucleotide sequences and corresponding amino acid sequences of the first linker peptide of selected sca genes of the present invention. $L_{ab}12$ (A), $L_{ab}14$ (B) and $L_{ab}15$ (C) respectively correspond to the constructions 1 (and 1A), 5 (and 5A) and 4 (and 4A) (with n=4) of Fig. 6. The sequences of the actual first linker peptides are underlined.

Detailed Description of the Invention

The single-chain antibody of the SCA of this invention is derived from a monoclonal antibody that is specific for a constituent of thrombi, preferably fibrin, particularly fibrin cross-links. Preferably,

the monoclonal antibody is a murine monoclonal antibody, such as MA-15C5, raised against human fibrin D-dimer. The properties of MA-15C5 have been described by Holvoet et al. (1989), and the construction of recombinant genes coding for the L- and H-chains of MA-15C5, from cDNA libraries of MA-15C5 hybridoma cells, has been described by Vandamme et al. (1990) and in European patent application ("EPA") 90401090.7. Parts of these recombinant genes can be used for the construction of an sca gene. In this regard, the amino acid sequences of the variable domains of the kappa and gamma chains of MA-15C5 are shown in Fig. 1 and Fig. 2, respectively. The CDR and framework regions are indicated in these Figures. The corresponding nucleotide sequences are shown in Fig. 3 and Fig. 4, respectively.

The SCA can have the following general structure NH_2-V_L-L_{ab}-V_H-COOH or NH_2-V_H-L_{ab}-V_L-COOH. In order to construct the SCA, the V_H and V_L domains should be linked by an appropriate first linker peptide (L_{ab}). A suitable L_{ab} can be designed using the computerized procedures outlined in PCT patent publication WO 88/01649 (which is incorporated herein by reference). Alternatively, the L_{ab} can be designed by the so-called "spare parts" method as described by Claessens et al. (1989). This method also involves the use, as a template, of an existing three-dimensional structure of an antibody molecule, with H- and L-chains similar to those of the fibrin-specific monoclonal antibody (e.g., MA-15C5) to be used for the construction of the SCA, to construct a 3D model of the fibrin-specific antibody or at least its framework regions. 3D structures of proteins can be found in, for example, the Brookhaven Database (Bernstein et al., 1977).

The design of the SCA also involves the following three steps:

1) Identifying suitable anchor regions in the V_L and V_H domains of the template antibody, to which the first linker peptide L_{sb} should be attached. The conformation of the anchor regions should be unaffected by introduction of the first linker peptide between them. The attachment sites at the ends of the anchor regions define a gap in which the linker sequence must be placed. The spatial distance between these attachment sites determines a minimum number of amino acids ("N_{aa}") that are necessary to bridge the gap. Parts of the V_L and V_H domains, flanking the anchor regions, can be considered as part of the first linker peptide.

2) Searching a database of 3D protein structures of sufficiently high resolution (e.g., lower than 3 Angstrom) to identify protein fragments, the ends of which overlap the anchor regions and which have the right length and 3D configuration to be able to serve as the first linker peptide, bridging the gap identified in step 1. As a general procedure fragments of proteins with a length of N_{aa} to N_{aa} + 6 are assessed as to their suitability as a first linker peptide. The ends of the fragments should fit the anchor regions (e.g., in a least square sense) as closely as possible so that the introduction of the fragment between the anchor regions will not change the correct association of the V_L and V_H domains. Identification of suitable fragments can be carried out by calculating the root mean square deviations ("rms") between the Cartesian coordinates of the alpha carbon atoms (or the main chain atoms) of the anchor regions and the overlapping regions of the protein fragments (Claessens et al., 1989). Only those fragments, for which the rms falls

below a certain threshhold determined by the user, are withheld for further study as an L_{ab}.

3) Selecting the most desirable fragments identified in step 2, which preferably conform to the following requirements:

- The first linker peptide should not interfere with the ordered secondary structure or with the folding of the V_L and V_H domains. Secondary structure predictions can be performed according to the procedures described by Jibrat et al. (1987).
- The regions of the first linker peptide that are exposed to solvent should not contain patches of hydrophobic residues.
- The first linker peptide should be sterically accomodated. Sterical accomodation of fragments can, for instance, be evaluated by calculating the non-bonded energy of the main chain atoms of the linker fragments with respect to the rest of the protein.

The amino acid sequences of the first linker peptide and/or the anchor regions can, if desired, be optimized by introduction of mutations (e.g., substitutions, deletions and/or additions) in order to reduce their non-bonded energy, to minimize their hydrophobicity and/or to improve their flexibility. In this respect, it may be preferred that regions of the first linker peptide, that immediately flank the anchor regions, be mutated to residues that were originally present in the V_L and V_H domains.

If no suitable anchor regions can be identified, for instance due to structural constraints or because the gap to be bridged is too big, the V_L and/or V_H can be extended by an extension sequence in appropriate

directions, and such a sequence can then serve as an anchor region.

For the construction of an SCA of this invention, it is preferred that the anchor regions be as near as possible to the appropriate C- and N- termini of the V_L and V_H or the V_H and V_L domains, respectively. However, an anchor region at the C-terminus of the V_H domain can also be located at the end of the V_H region that is encoded by the J gene segment.

3D structures of proteins (e.g., SCAs) of this invention can be obtained by methods such as crystallography (Wyckoff et al., 1985), nuclear magnetic resonance spectroscopy (Wüthrich, 1986), structure derivations based on available 3D structures from homologous proteins (see, e.g., Blundell et al., 1987), or from structure predictions based on analysis of the primary structures (for a review, see Taylor, 1988).

The 3D structures of proteins of this invention can be analyzed and modelled by the use of a dedicated computer software package such as the BRUGEL (R) molecular graphics software package (Delhaise et al., 1985 - Plant Genetic Systems N.V., Ghent, Belgium). The effects of substitutions, deletions and additions in known 3D structures or template-derived models on the conformation of the proteins can also be so-analyzed.

In accordance with a preferred embodiment of this invention, the SCA is characterized by an amino acid sequence as shown in Fig. 6. Alternatively, use can be made in such an SCA of a first linker peptide which comes from a naturally-occurring protein and which seems to serve as a natural linker between major functional domains of the naturally-occurring protein, such as a hinge-like sequence of an immunoglobulin.

For the production of a SCAPA of this invention, the corresponding SCA can be connected to at least the catalytic domain of a plasminogen activator (i.e., the PA-portion). For this purpose, the serine-protease catalytic domain of scuPA (comprising amino acids 144 to 411 in Fig. 5) is preferred, but the catalytic domains of tPA and other plasminogen activators can also be used. The SCA- and PA-portions can, in principle, be connected in two ways: the C-terminus of the PA-portion can be linked to N-terminus of the SCA or the C-terminus of the SCA can be linked to the N-terminus of the PA-portion. In order to link the PA-portion to the SCA-portion, a suitable second linker peptide (L_{cd}) should be designed. If the 3D structures of plasminogen activators of interest (or proteins with appreciable homology thereto) are available, this can be done using the same procedure as is used for the construction of the first linker peptide between V_L and V_H (or vice versa).

If 3D structures are not available for the second linker peptide, use can be made of sequences which seem to serve as natural linkers between major functional domains of the PA-portion. For instance, folding experiments have shown that, in scuPA, the region between Ala132 and Leu144 serves as a natural linker between the kringle and catalytic domains (Oswald et al., 1989). Consequently, when the catalytic domain of scuPA is used for the production of the SCAPA of this invention, in which the C-terminal part of a SCA-portion is linked to the N-terminal part of a PA-portion, it is preferred that part or all of this region be used as a second linker peptide between the two portions. In such a case, preferred attachment sites on the scuPA are believed to be Ala132, Lys136 and Leu144. In all of these cases, preferred attachment

sites at the C-terminal part of the V_L-L_{ab}-V_H of SCAs are believed to be Ser113 (i.e., the end of the V_H part encoded by the J gene segment), Ser120 (i.e., the actual end of the fourth framework of the V_H domain), or any other amino acid between these two residues of the heavy chain. A preferred C-terminal attachment site on the V_H-L_{ab}-V_L of SCAs is believed to be located at Leu104 of the kappa-chain (numbering as in Fig. 1 and Fig. 2).

Alternatively, part of the constant domain following the variable domains of the heavy and light chains of the fibrin-specific antibodies can also be used as the second linker peptide between the SCA- and PA-portions. As another alternative, all or part of the A-domain of scuPA, such as the EGF-like and/or the Kringle domains of scuPA, can be used as the second linker peptide.

It is preferred that proteolytic cleavage of the uPA catalytic domain (e.g., by plasmin or thrombin), resulting in inactive forms of the protein, be prevented. This can be most conveniently done by mutating the amino acids at the cleavage sites so that they are no longer recognized by the proteolytic enzymes. In this regard, Phe157 of scuPA can, for instance, be mutated to Asp157 to remove the Arg156-Phe157 thrombin cleavage site. If also needed, Lys135 of scuPA can be mutated, for instance to Gln135, to remove the Lys135-Gln136 plasmin cleavage site.

The SCA and the SCAPA of this invention can be produced by the expression, in host cells, of the sca and scapa genes, respectively, preferably the chimaeric gene of this invention. The construction of these genes can be achieved in a conventional manner. cDNAs coding for V_L and V_H can, for instance, be isolated from cDNA libraries from suitable hybridomas producing thrombus-

specific, preferably fibrin-specific, antibodies (see, e.g., Vandamme et al, 1990). DNA fragments coding for linkers L_{ab} and L_{cd} can be directly synthesized. The gene coding for prourokinase can be obtained as described by Holmes et al. (1985). Appropriate DNA fragments can be ligated to each other by conventional means so as to produce one contiguous DNA fragment coding for the SCA or SCAPA protein of this invention.

The sca and scapa genes can be expressed in suitable procaryotic or eucaryotic host cells by placing the genes under the control of a promoter capable of directing their expression in the host cells. Conventional promoters can be used. Preferred promoters for use in E. coli are, for example, regulatable promoters such as: the P_{tac} promoter (De Boer et al., 1983), the sequence of which is shown in Fig. 7, the P_{lac} promoter (Fuller, 1982), the P_{trp} promoter (Martial et al., 1979), the lambda P_L promoter (Bernard et al., 1979) and the P_R promoter (Zabeau and Stanley, 1982). Preferred promoters for use in mammalian cells have, for example, been described by Menck et al. (1987), Baker et al. (1988), Artelt et al. (1988) and Lee et al. (1981).

If required, a signal sequence can be placed in front of, and in reading phase with, the sca or scapa gene. The signal sequence provides: a) a translation initiation site and b) the necessary functional sequence for exporting the SCA or SCAPA. By signal sequence is meant a DNA fragment coding for a polypeptide fragment ("signal peptide") which is normally associated with a protein, or subunit of a protein that is translocated out of the cytosol of the host cell --for example, to the periplasmic space in E. coli, to the medium in B.subtilis, or to the endoplasmic reticulum (and, if no other targeting

information is available, to the medium or extracellular space) in eucaryotic cells (or tissues) derived from organisms such as yeasts, insects or mammals. The signal peptide is responsible for the translocation process during which the signal peptide is separated or proteolytically removed from the protein or subunit. Signal sequences which can be used are those coding for the signal peptides listed by Watson (1984) or for signal peptides that conform to the general characteristics as outlined by Von Heyne (1988). A preferred signal sequence that can be used in E. coli is the one coding for the phoA signal peptide (Michaelis et al., 1983) which is shown in Fig. 7. Preferred signal sequences, that can be used in eucaryotic cells, are those coding for the signal peptides that are naturally associated with the heavy and light chains of antibodies. In this regard, the amino acid sequence and its encoding nucleotide sequence of the signal peptide of the kappa chain of the MA-15C5 monoclonal antibody is shown in Fig. 8A, and the amino acid sequence and its encoding nucleotide sequence of the signal peptide of the gamma chain of MA-15C5 is shown in Fig. 8B. Another preferred signal sequence is that coding for the signal peptide normally associated with the plasminogen activator, for instance the signal peptide associated with scuPA (Fig. 5 - residues Met1 to Gly20 - see also Holmes et al., 1985).

Although the use of a signal peptide is preferred for the production of the SCA or SCAPA, it is not necessary. A cell can be transformed with just a sca or scapa gene encoding a SCA or SCAPA under the control of a suitable promoter, and the SCA or SCAPA, expressed by the transformed cell intracellularly, can be obtained by lysing the cell.

Preferred host cells to express the chimaeric gene of this invention are insect cells. In this respect, use can be made of the baculovirus expression vectors (see Maeda, 1989 for a general review). Preferably, the chimaeric gene is placed under the control of the strong polyhedrine promoter, particularly of the Autographa californica nuclear polyhedrosis virus and expressed in Spodoptera frugiperda using the procedures and vectors described, for example, by Summers and Smith (1987) and Luckow and Summers (1987, 1989) and in US patent 4745051. Other baculovirus expression vectors, such as those described in EP 345152 and EP 340359 and PCT publications WO 89/01038 and WO 89/01037, can also be used.

The SCA and SCAPA can also be prepared by construction of a chimaeric gene capable of being expressed in other host cells, such as E. coli, B.subtilis, yeasts and mammalian cells (e.g., CHO cells), preferably mammalian cells. Appropriate promoters, regulatory sequences (including 3' regulatory sequences, as well 5' regulatory sequences and enhancer sequences) and, if necessary, signal sequences for such a chimaeric gene are well-known to those skilled in the art.

The chimaeric gene of this invention can be introduced into host cells, the host cells can be cultured, and the SCA or SCAPA can be purified from the host cell culture by conventional means. Secreted SCA or SCAPA can be purified, for example, by affinity chromatography on immobilized epitope (e.g. D-dimer) and/or immunoadsorption to insolubilized antibody raised against the PA-portion.

It goes without saying that the SCA of this invention can be constructed using the variable domain of thrombus-specific, preferably fibrin-specific,

monoclonal antibodies of other than MA-15C5. The corresponding sca gene can be constructed and expressed in analogous ways to those described above. Monoclonal antibodies that can be used are, for instance, those described by Kudryk et al. (1984), Elms et al. (1983), Scheefers-Borchel et al. (1985) and Hui et al. (1986) and in Australian patent publication AU-B-25387/84.

When the SCA or SCAPA of this invention is used for multiple intravenous applications in patients, it may be preferred to minimize its immunogenicity. This can be achieved (see, e.g., LoBuglio et al., 1989) by replacing the nucleotide sequences coding for the murine framework regions by the corresponding sequences coding for framework regions derived from variable domains of human antibodies as described, for example, by Riechmann et al. (1988) and Verhoeyen et al. (1988) and in EP 328404.

The SCA of this invention can be used for imaging of thrombi. The SCA can be labelled with an opacifying agent, such as an NMR or X-ray contrasting agent, or radioactively labelled in a conventional manner.

The SCAPA of this invention can be used as a thrombolytic agent to treat patients with myocardial infarction, peripheral arterial thrombosis, and stroke, as well as deep venous thrombosis and pulmonary embolism. The SCAPA has a number of advantages over existing thrombolytics. The thrombus-specific SCA-portion targets the SCAPA, and consequently its plasminogen activation activity, to the thrombus. The use of an SCA derived from an antibody specific for fibrin, particularly fibrin cross-links (such as MA-15C5), is especially preferred. This ensures that the corresponding SCAPA will remain in contact with the thrombus for a longer period of time during the degradation of fibrin. Because it is believed that the

half-life of the SCAPA is likely to be predominantly determined by its SCA-portion (Collen et al., 1989), it is also expected that the half-life of such a molecule will be greater than its PA-portion alone. It is believed that the half-life of the SCAPA of this invention also can be increased by producing it: 1) in a non-glycosylated form or in a super-glycosylated form or in a form in which some glycosylation is added to the SCAPA (i.e., to one or more regions of the SCAPA) and other glycosylation is removed from the SCAPA (i.e., from one or more other regions of the SCAPA); and/or 2) in a form which is resistant to plasminogen activator inhibitors; and/or 3) with all or at least a significant part of the A domain (at the N' end of the catalytic domain) of its PA-portion, particularly of scuPA, serving as the second linker peptide. This will permit the application of the SCAPA as a bolus injection and will possibly result in a reduction of the incidence of reocclusion. When compared to other combinations of fibrin-specific antibodies and thrombolytic agents, it is also expected that the SCAPA of this invention will display a lower immunogenicity and a better thrombus penetration due to its reduced molecular weight.

The final conformation of the SCAs and SCAPAs of this invention will depend upon the independent folding of their separate domains (V_L, V_H and plasminogen activator catalytic domain) and not upon the association of disulfide bridges with separate polypeptide chains. It is believed that this will simplify production of these molecules in host cells transformed with sca or scapa genes of this invention as described above. In this regard, it believed that transformed insect or mammalian host cells can properly process and secrete the SCAs and SCAPAs of this

invention, so that they are properly folded for binding to a thrombus constituent and without significant loss of binding activity (as compared to the thrombus-specific antibodies, from which they are derived). Furthermore, the SCAs and SCAPAs of this invention can, if desired, be produced in a glycosylated or super-glycosylated form in insect or mammalian cells transformed with, respectively, sca or scapa genes or mutated sca or scapa genes in which glycosylation sites have been added. Alternatively, the SCAs and SCAPAs can be conveniently expressed in transformed host cells in a non-glycosylated form by mutating the sca and scapa genes at sites which would otherwise encode amino acid sequences which could be glycosylated. In this regard, potential glycosylation sites could be eliminated, for example: in the portion of the nucleotide sequence of Figure 4 encoding the heavy chain domain (V_H) of the MA-15C5 antibody, by mutating its AAT nucleotides encoding Asn at positions 261-263 to the nucleotides GAT encoding Asp; and/or in the portion of the nucleotide sequence of Figure 5 encoding scuPA, by mutating its AAT nucleotides at positions 1063-1065 to the nucleotides GAT. In addition, the SCAPAs of this invention can be conveniently expressed in transformed host cells in a form more resistant to a plasminogen activator inhibitor (e.g., PAI-1) by mutating the plasminogen activator catalytic domain encoded by the scapa gene, so that it encodes, for example, a mutant tPA-encoding region as described by Madison et al. (1989, 1990) or a mutant SCUPA-encoding region in which, from nucleotide 691 to nucleotide 702 in Figure 5, the amino acids Arg Arg His Arg are changed to smaller uncharged amino acids such as Ala or to negatively charged amino acids such as Glu.

The Examples, which follow, illustrate this invention. Unless otherwise stated in the Examples, all procedures for making and manipulating recombinant DNA were carried out by the standardized procedures described in Sambrook et al, "Molecular Cloning - A Laboratory Manual", Cold Spring Harbor Laboratory (1989). All modelling and analysis of 3D structures of proteins was performed using the BRUGEL(R) software package (Plant Genetic Systems N.V., Ghent, Belgium). All mutagenesis was performed by oligonucleotide-directed construction of mutations by the gapped duplex DNA method (Kramer et al., 1984) using the pMa/c vectors described by Stanssens et al (1987, 1989). Appropriate oligonucleotides were designed according to the general rules outlined by Kramer and Fritz (1988) and synthesized by the phosphoramidite method (Beaucage and Caruthers, 1981) on an Applied Biosystems 380A DNA synthesizer (Applied Bisosystems B.V., Maarssen, Holland).

Example 1 : Design of first linker peptides for the V_L-L_{ab}-V_H and V_H-L_{ab}-V_L SCAs

In this and the following examples, the numbering of residues of the MA-15C5 V_L and V_H domains will follow the standardized numbering of Kabat et al. (1987) (see Figures 1 and 2). The numbering of residues of scuPA will follow that of Holmes et al. (1985) and that of Fig. 5. The numbering of residues of proteins for which the actual 3D structure is known and available in a public database, such as the Brookhaven Database (Bernstein et al., 1977), will follow the numbering as used in this database.

1. Modelling of the 3D structure of the MA-15C5 monoclonal antibody.

The MA-15C5 antibody contains a kappa light chain and a gamma heavy chain. Thus, the Brookhaven Database was searched for structures of immunoglobulines with similar heavy and light chains. The protein with code pdb2hfl, which is a Fab-lysozyme complex (Sheriff et al., 1987), fulfilled these requirements.

A model of the MA-15C5 V_L and V_H was obtained by substitution of all residues of the pdb2hfl structure that differed from the MA-15C5 V_L and V_H sequences with their corresponding residues in the MA-15C V_L and V_H.

All substitutions were carried out in the absence of explicit hydrogen (i.e., the no-hydrogen model) and in the absence of water molecules and sequentially from the N- to C-terminus of the V_L and V_H domains. Main chain atoms were taken from pdb2hfl template. Side chain orientations were determined by exhaustive map computation varying each of the side chain dihedral angles in steps of 30° in the 0-360 interval and by selecting the configuration with the lowest energy.

The deletions and insertions that would normally be required were not introduced because it was observed that the locations of these mutations were not in regions (i.e., the framework regions) that were important with respect to the linker construction.

2. Construction of V_L-L_{ab}-V_H.

While looking for a suitable anchor region at the N-terminus of the MA15C5 V_H, it was observed that the first two residues of the gamma-chain had a high temperature factor, thus reflecting mobility. Therefore, these residues were not included in the anchor region. Thus, the anchor region of V_H was defined as the segment comprising residues 3 to 7 (i.e., QLKQS) which forms the end of a β-sheet.

A suitable anchor region at the C-terminus of the MA-15C5 V_L was found to be the segment comprising

residues 102 to 106 (i.e., TKLEI) which also forms the end of a β-sheet.

The gap between the attachment points of the two linkers is 30.8 Angstrom. Thus, a first linker peptide (L_{ab}) of at least 8 amino acids should be sufficient to bridge the gap. Note that there are still two residues (i.e., KR) flanking the C-terminus of the V_L anchor region, and two residues (i.e., QV) that flank the N-terminus of the V_H anchor region.

Proteins in the Brookhaven database, that are refined and have a resolution lower than 3 Angström, were searched for fragments which consisted of terminal regions overlapping the anchor regions of V_L and V_H and central regions capable of bridging the gap between the attachment points of the anchor regions. The fitting of the fragment terminal regions with the V_L and V_H anchor regions was assessed by a least square fit of the atomic coordinates of: 1) the alpha carbon atoms and 2) all main chain atoms (MacLachlan, 1979). This analysis resulted in a root mean square deviation (rms) which should be minimal. The best fragment was found to be the segment comprising residues 22-42 from proteinase K (pdb2prk - Betzel et al., 1988). The following alignment could be made (the anchor regions or the L_{ab} sequence between them are underlined):

...TKLEIKR QVQLKQS... (V_L gap V_H)

TYYYDESAGQGSCVYVIDTGI (pdb2prk fragment)

The anchor regions of V_L and V_H were, of course, retained. The structure was then subjected to 100 steps of a "Steepest Descent" (Fletcher and Reeves, 1964) energy minimalization procedure fixing all atoms except those of the first linker peptide.

The residues of the first linker peptide that overlap with other regions of V_L and V_H were mutated to the residues that were originally present in the V_L and

V_H. Consequently, the real first linker peptide between the V_L and V_H chains was only 7 residues long (i.e., AGQGSCV). The C residue in this first linker peptide was mutated into a S residue to prevent unwanted disulfide bridge formation. The non-bonded energy of this linker with respect to the rest of the protein was good ($E_{Van\ der\ Waals}$=-39 kcal, E_{elec}=-8.5 kcal--both calculated in the no-hydrogen model). The energy of the overall structure was also favorable ($E_{Van\ der\ Waals}$=-1355 kcal, E_{elec}=-190 kcal--no-hydrogen model). No hydrogen bonds were observed to be formed between the linker and the rest of the structure, and no cavities were created. The final SCA is shown in Fig. 6 (constructions 1 and 1A).

The first linker peptide was observed to be located at the side opposite to the antigen-binding site and should not interfere with binding. It was also seen that mutation of the Ile106 and Arg108 residues into Gly or Ser residues also resulted in suitable SCAs (Figure 6, constructions 2, 3, 2A and 3A). Flexibility of the first linker peptide could be increased by replacement of the Q residue in the linker with R followed by 0 to 4 glycine residues (Fig. 6, constructions 4 and 4A) or by replacing the AGQ block of residues in the linker by one or more GGGS blocks of residues.

It was also attempted to use the linker proposed by Bird et al. (1988). This linker has the sequence KESGSVSSEQLAQFRSLD. It was found that the most favorable construction was that in which this linker was attached to Leu104 of V_L and Val2 of V_H (Fig. 6 constructions 5 and 5A).

3. Construction of V_H-L_{ab}-V_L.

As the N-terminus of V_L is located near the CDR region, the first linker peptide should satisfy the

stringent requirement of not interfering with the antigen binding sites. The N-terminus is located at the start of a β-strand situated at the edge of one of two antiparallel β-sheets packed on top of each other. Interference with the antigen-binding site can be prevented by designing an extension of the V_L chain so that the β-sheet at the N-terminus is entered by a β-turn preceded by a β-strand. An additional advantage of such an extension is that the gap between the C-terminus of the V_H and the N-terminus of the (extended) V_L is smaller so that a shorter first linker peptide is required. As immunoglobulins contain a large collection of turn motifs, the Fab structure (pdb2hfl--Sheriff et al., 1987) was searched for turns for which the main chain heavy atoms of their C-terminal ends could be fitted (in a least square sense) on the main chain heavy atoms of the first three N-terminal amino acids of V_L. It was observed that the best fit could be obtained by fitting the segment of amino acid residues at positions 68 to 70 of pdb2hFL on the first three amino acids of the V_L of MA-15C5 (rms = 0.67 Angstrom). The resulting configuration can be schematically represented as follows (the anchor region and extension are underlined):

.....SVTVSS DIKM... (V_H gap V_L)

SGSGSGTSY (pdb2hfl fragment)

Residues 68 to 70 of pdb2hfl are preceded by a β-strand segment (residues 62 to 67). Upon fitting the 68-70 residues of the 62-70 2HFL fragment (i.e. SGSGSGTSY) on the first three residues of the MA-15C5 V_L it was observed that the β-sheet of MA-15C5 was extended by one β-strand.

The actual first linker peptide was then designed between the C-terminus of the V_H domain (using the residues 108-111, i.e., SVTV, as an anchor region) and

the N-terminus of the extended V_L domain using the first four residues of the V_L extension (i.e., SGSG) as an anchor. A search of the 3D structures of proteins for suitable fragments resulted in the identification of a fragment from pdb2sod (superoxide dismutase ("SOD") -- Tainer et al., 1982) with 11 residues (SOD residues 038-048 - EGDHGFHVHQF) between the anchor regions. The configuration of the fit can be represented as follows (the anchor regions and first linker peptide are underlined):

...SVTVSS SGSGSGDIKM... (V_H gap extension-V_L)
TGLTEGDHGFHVHQFGDNT (pdb2sod)

The total first linker peptide (EGDHGFHVHQFSGSGSG) between the original V_H and V_L domains is thus composed of this 11 residue pdb2sod fragment plus the six residue pdb2hfl extension sequence which was introduced at the N-terminus of the V_L domain.

The structure was then subjected to 100 steps of a "Steepest Descent" (Fletcher and Reeves, 1964) energy minimalization procedure fixing all atoms except those of the first linker peptide. The first two amino acids of the pdb2sod linker fragment (EG) were then mutated to serine residues to revert to the original V_H C-terminus. Furthermore, the linker's hydrophobicity was reduced by mutating:

- the F residue at linker position 6 into S
- the V residue at linker position 8 into S
- the F residue at linker position 11 into S

In addition, the S residue at linker position 12 was mutated into a G to increase flexibility. The H residue at linker position 4 was initially not mutated because it was observed that the imidazole was involved in hygrogen bonding with groups in V_H.

The final construction (see also Fig. 6, construction 6) was thus :

{...SVTVSS}-(DHGSHSHQS-GGSGSG)-[DIKM...]
corresponding to {V_H}-(linker)-[V_L].

Further modulations of the polarity and hydrophilicity of the first linker peptide can be made by additional modification to it, resulting in the following constructions :

{...SVTVSS}-(DHGSHSEQSGGSGSG)-[DIKM...]
{...SVTVSS}-(GGGSHSEQSGGSGSG)-[DIKM...]
{...SVTVSS}-(GGGSGSGGSGGSGSG)-[DIKM...]
{...SVTVSS}-(GGGGSGGGGSGGGGS)-[DIKM...]

(Fig. 6, constructions 7 to 10). The last of these first linker peptides corresponds to the linker that was used by Huston et al. (1988).

Example 2 : Construction of sca genes and baculovirus expression vectors containing these genes

The PstI-HindIII fragment of Fig. 4 contains most of the V_H domain and part of the N-terminal part of the C_{H1} domain of the MA-15C5 gamma chain. Only the first four amino acids of the V_H domain are not present (Gln-Val-Gln-Leu). This fragment was cloned into the PstI and HindIII sites of pUC19 (Yanisch-Perron et al., 1985). The SmaI-Hind III fragment of the resulting plasmid, pUC19-gamma6 was then cloned in pMc5-8-uts digested with EcoRI, filled in with the Klenow fragment of E. coli DNA polymerase I (Klenow), and further digested with HindIII, yielding plasmid pMc5-gamma6-S. Plasmid pMc5-8-uts can be obtained by cloning a universal translation stop sequence ("uts") with the following sequence :

AGCTTGCTGATTGATTGACCGGATCGATCCGGCT
ACGACTAACTAACTGGCCTAGCTAGGCCGAGATC

between the HindIII and XbaI sites of the polylinker of pMc5-8 which was described by Stanssens et al. (1987, 1989).

pMc5-gamma6-S can be used directly for site directed mutagenesis. A stop codon and a EcoRI site was introduced immediately after Ser113 by introduction of the sequence TGAATTC, yielding pMc5-G60-S. The EcoRI sites, and the sequences between them were then deleted by digestion of pMc5-G60-S with EcoRI (filled in with Klenow) and religation. The resulting plasmid was designated as pMc5-G60ΔE-S.

The kappa chain was obtained on plasmid pCMBDHFR13-15C5KMu (Vandamme et al., 1990). The EcoRI-BglII fragment, shown in Fig. 3, was cloned in the EcoRI and BamHI sites of pMc5-8, yielding plasmid pMc5-Kb. This fragment comprises the signal peptide, the V_L domain and the C_L domain of the Ma-15C5 kappa chain. The EcoRI (filled in with Klenow)-XbaI fragment of pMc5-kb, comprising the kappa chain, was then cloned in the BamHI (filled in with Klenow) and XbaI sites of the baculovirus expression vector pVL1393, yielding pVL1393-K. pVL1393 (now available from British Biotechnology Ltd., Oxford, UK) can be obtained from pVL941, described by Luckow and Summers (1989), by deletion of a 630 bp EcoRI-XmaIII fragment and by extension of the polylinker by insertion of the following sequence in the BamHI site of the pVL941 polylinker :

GATCCCGGGTACCTTCTAGAATTCCGGAGCGGCCGCTGCAGATCT

GGCCCATGGAAGATCTTAAGGCCTCGCCGGCGACGTCTAGACTAG

(Summers, personal communication).

The construction of the sca gene coding for the SCA was done as follows. pMc5-G60ΔE-S was digested with AccI (filled in with Klenow) and XbaI and the fragment, containing the V_H encoding sequence, was cloned in the StyI (filled in with Klenow) and XbaI sites of pMc5-kb, yielding plasmid pMc5-KG60ΔE-S. In one mutagenesis experiment, appropriate transition sequences between

the V_L and V_H coding regions were then provided. This involved deletion of the remaining part of the kappa constant region, appropriate mutagenesis of the C-terminus of the V_L and the N-terminus of the V_H domains (including the addition of the missing N-terminal amino acids of V_H), and addition of a linker sequence. Three such mutagenesis experiments, with different linker sequences, led to folowing three plasmids : pMc5-K12A, pMc5-K14A, pMc5-K15A5 (corresponding to SCAs 1, 5 and 4 respectively in Example 1). These plasmids differ by the first linker peptide between the V_L and V_H domains which are shown in Fig. 9.

The sca genes in pMc5-K12A, pMc5-K14A, pMc5-K15A5 were then introduced in pVL1393-K by replacement of the BamHI-XbaI fragment (coding for the C-terminal end of the MA-15C5 kappa chain) of pVL1393-K with the BamHI-XbaI fragments of pMc5-K12A, pMc5-K14A, pMc5-K15A5, yielding plasmids pVL-K12A, pVL-K14A, pVL-K15A5. These plasmids can be used directly for transfection of insect cells.

Example 3 : Construction of scapa genes and baculovirus expression vectors containing these genes

The 1475 bp HindIII fragment of the scupa cDNA (Fig. 5) was cloned in the HindIII site of pUC18, yielding plasmid pULscu-PA (Nelles et al., 1987). The NcoI (filled in with Klenow)-HindIII fragment of this plasmid was further subcloned in the BamHI (Klenow) and HindIII sites of pMC5-8, yielding plasmid pMc5-scupa-Nco.

In four subsequent mutagenesis steps carried out on plasmid pMc5-scupa-Nco, the C at position 1356 (Fig. 5) was mutated to a T (resulting in a destruction of a BamHI site), the G at position 966 (Fig. 5) was mutated to an A (resulting in a destruction of a FspI site),

the AAA codon at position 562 (Fig. 5), coding for Lys-135 in scuPA, was mutated to a CAA codon (Gln), and the TTT codon at positions 628-630 (Fig. 5), coding for Phe157 of scuPA, was mutated to a GAT codon (Asp). The resulting plasmid was designated as pMc5-scupa-77-I.

In five additional mutagenesis steps carried out on plasmid pMc5-scupa-77-I, the A at position 648 (Fig. 5) was mutated to a G (resulting in a destruction of an EcoRI site), the G at position 1092 (Fig. 5) was mutated to an A (resulting in the destruction of a PvuII site), the AGGs at positions 691-696 (Fig. 5) were each mutated to a CGT, the G at position 702 (Fig. 5) was mutated to a C (resulting in the creation of a SacII site), and the C at position 624 (Fig. 5) was mutated to a T (resulting in the creation of an StuI site). The resulting plasmid was designated as mPc5-scupa-77-II.

The FspI-XbaI fragments of the pMc5-scupa-77-I and -II plasmids (each comprising the coding sequence of the scuPA catalytic domain) were cloned into the HindIII (filled in with Klenow) and XbaI sites of pMc5-K12A, pMc5-K14A, pMc5-K15A5 (from Example 2), after which the sequences between the C-terminal Ser113 (Fig. 2) codon of the Ma-15C5 V_H and the N-terminal Ala132 codon of the scuPA were deleted. The resulting plasmids were designated as pMc5-K12A-PA-I, pMc5-K14A-PA-I, pMc5-K15A5-PA-I, pMc5-K12A-PA-II, pMc5-K14A-PA-II and pMc5-K15A5-PA-II, respectively.

The scapa genes in the pMc5-K12A-PA-I and -II, pMc5-K14A-PA-I and -II, and pMc5-K15A5-PA-I and -II plasmids were then introduced in the transfection vector pVL1393-K (from Example 2) by replacement of a BamHI-XbaI fragment (coding for the C-terminal part of the MA-15C5 kappa chain) of pVL1393-K with the BamHI-XbaI fragments of pMc5-K12A-PA-I and -II, pMc5-K14A-

PA-I and -II, and pMc5-K15A5-PA-I and -II, yielding plasmids pVL-K12A-PA-I, pVL-K14A-PA-I, pVL-K15A5-PA-I, pVL-K12A-PA-II, pVL-K14A-PA-II and pVL-K15AS-PAII, respectively. These plasmids can be used directly for transfection of insect cells.

Example 4 : Expression of sca and scapa genes in insect cells

The sca and scapa genes of Examples 2 and 3 (in plasmids pVL-K12A, pVL-K14A, pVL-K15A5, pVL-K12A-PA-I and -II, pVL-K14A-PA-I and -II and pVL-K15A5-PA-I and II) are introduced and expressed in Spodoptera frugiperda (SF9) cells (ATCC no. CRL 1711) using the procedures and Autographa californica nuclear polyhedrosis viruses (AcNPV) described by Summers and Smith (1987).

Example 5 : Purification of SCAs and SCAPAs from insect cell cultures of Example 4

The secreted SCAs expressed in Example 4 are purified by means of affinity chromatography on immobilized fibrin fragment D-dimer.

For purification of secreted SCAPAs expressed in Example 4, this step is followed by immunoadsorption on an insolubilized monoclonal antibody against urokinase, MA-4D1E8, as described by Nelles et al. (1987). The fractions containing urokinase-related antigen are pooled and dialyzed against 0.3 M NaCl, 0.2 M arginine, 0.02 M Tris.HCl buffer pH 7.4, containing 0.01% Tween 80 and 10 KIU/ml aprotinin.

tcuPA is removed from samples equilibrated with dialysis buffer containing 0.2 M arginine by chromatography on benzamidine-sepharose. Fractions devoid of amidolytic activity are pooled.

Aprotinin is removed by extensive washing on a Centrocon 30 microconcentrator (from Amicon, Danvers, MA, USA).

Example 6 : Purification of SCAPAs from insect cultures of Example 4

The secreted SCAPAs expressed in Example 4 are also purified in a different way from that of Example 5. Each SCAPA is purified by ion exchange chromatography on SP-Sephadex (from LKB, Bromma, Sweden), followed by gel filtration on Sephadex-G100 superfine (from LKB).

In the first step, 1.5 l of conditioned medium, with a pH adjusted to 5.5, is applied at 4°C and a flow rate of 20 ml/h on a 0.9 x 2 cm SP-Sephadex column equilibrated with 0.05 M NaH_2P04, pH 5.5, containing 0.05 M NaCl, 0.01% Tween 80 and 10 KIU/ml aprotinin. Elution is performed with a 60 ml gradient from 0.05 M to 0.60 M NaCl in 0.05M NaH_2PO_4, pH 5.5. The fractions containing each SCAPA, as determined with an ELISA specific for uPA-related antigen, are pooled, and the pH is increased to 7.4 with 1 M NaOH. The pooled fractions (representing 7 ml with a concentration of 0.28 mg of SCAPA per ml) are concentrated on a Centricon 30 microconcentrator (Amicon) to a final volume of 0.5 ml. The concentrated sample is then applied at 4°C and at a flow rate of 4 ml/hr on a 1.0 x 110 cm Sephadex-G100 superfine column equilibrated with 0.02 M Tris-HCl buffer, pH 7.4, containing 0.3 M NaCl, 0.01% Tween 80 and 10 KIU/ml aprotinin. The fractions containing the SCAPA are pooled. Aprotinin is then removed by extensive washing on a Centricon 30 microconcentrator with 0.05 M Tris-HCl buffer, pH 7.4, containing 0.038 M NaCl and 0.01% Tween 80.

Example 7 : Evaluation of the purified SCAs and SCAPAs from Examples 5 and 6.

Satisfactory antigen-binding activity of the SCAs and SCAPAs of Examples 5 and 6 is found in ELISA using immobilized fibrin fragment D-dimer and rabbit-anti-mouse antibodies specific for MA-15C5 and goat antibodies specific for total rabbit IgG fraction conjugated to alkaline phosphatase (Voller et al., 1976). Satisfactory urokinase-related antigen activity is also found in ELISA according to Darras et al. (1986).

The SCAs and SCAPAs are also characterized by SDS-PAGE under reducing and non-reducing conditions, and the amino termini of the proteins are determined to verify correct processing. Satisfactory equilibrium association constants of the SCAs for immobilized and dissolved purified fragment D-dimer are determined according to Hogg et al. (1987). SCA is labeled by ^{125}I to show that there is satisfactory in vitro plasma clot binding capacity (Lijnen et al., 1986) and to determine in vivo half-life.

Satisfactory specific activity of the SCAPAs is shown on fibrin plates (Astrup et al., 1952) by comparison with the International Reference Preparation for urokinase (Nelles et al., 1987). The SCAPAs are treated with plasmin (Lijnen et al., 1988) to produce two-chain SCA-tcuPA variants, after which satisfactory amidolytic activity on the synthetic substrate pyroglutamyl-glycyl-arginine-pNA (S-2444 - Kabi-Diagnostica) is shown. Satisfactory plasminogen activation activity of the SCAPAs is measured in the presence of an excess of the synthetic substrate S-2251 (Kabi-Diagnostica). Satisfactory in vivo plasma clot binding capacity for ^{125}I-labeled SCAs is shown in a rabbit jugular vein thrombolysis model (Collen et al.,

1983), and satisfactory in vivo plasma clot lysis activity of the SCAPAs is shown in the quantitative rabbit jugular vein thrombolysis model (Collen et al., 1983) and in the quantitative dog arterial thrombolysis model (Yasuda et al., 1989), in the dog coronary thrombolysis model (Bergman et al., 1983), and in the baboon coronary thrombolysis model (Flameng et al., 1985).

Example 8 : Synthesis and Expression of a pVL-K12A' as in Examples 1 and 2

The computer assisted method of Claessens et al. (1989) was used for the design of a synthetic linker L12 for connecting the carboxyterminal end of the V_L domain of MA-15C5 to the aminoterminal end of its V_H domain. Since the variable domains of antibodies appear to have homologous three-dimensional structures, modeling was based on pdb2hfl (Sheriff et al., 1987). Design of the polypeptide linker was initiated by selecting anchor amino acids (i.e., residues with low temperature factor, reflecting low mobility). A suitable anchor region at the carboxyterminus of the V_L domain of MA-15C5 was found to be the segment comprising residues Thr^{102}-Lys^{103}-Leu^{104}-Glu^{105}-Ile^{106}, a segment that is at the end of a β-sheet. A suitable anchor region at the aminoterminus of the V_H domain of MA-15C5 was found to be the segment comprising residues Gln^3-Leu^4-Lys^5-Gln^6-Ser^7, a segment that also is at the end of a β-sheet. The attachment sites at the ends of the anchor regions define a gap in which the linker has to be fitted. The 30.8 Å spatial distance between these attachment sites determines a minimum number of amino acids that are required to bridge the gap. This minimum number was found to be 8. The Brookhaven Protein Database was then searched for all peptide sequences

consisting of 16 to 22 amino acids, so that the length of the peptide linker could be varied from 8 to 14 amino acids. To overlap the V_L carboxyterminal anchor region, 5 more amino acids were required. To overlap the V_H aminoterminal anchor region, 3 more amino acids were required. Thus, the number of amino acids had to vary between 16 and 22. This search yielded more than 10,000 peptide candidates. Secondary structure predictions were then performed according to Jibrat et al. (1987) to select those peptide segments that did not interfere with the ordered secondary structure or with the folding of the V_L or the V_H domain of MA-15C5. In this way, the number of linker peptide candidates was reduced to 82.

The fitting of the peptide terminal regions with the V_L and V_H anchor regions was assessed by a least square fit of atomic coordinates of the α carbon atoms and of all main chain atoms (MacLachlan, 1979). This analysis resulted in a root mean square deviation that was minimal for a 20 amino acid sequence: Thr-Tyr-Tyr-Tyr-Asp-Glu-Ser-Ala-Gly-Gln-Gly-Ser-Cys-Val-Tyr-Val-Ile-Asp-Thr-Gly-Ile, derived from proteinase K (Betzel et al., 1988). In this sequence, the Thr-Tyr-Tyr-Tyr-Asp fragment overlapped the Thr^{102}-Lys^{103}-Leu^{104}-Glu^{105}-Ile^{106} V_L carboxyterminal anchor region, and the Ile-Asp-Thr-Gly fragment overlapped the Gln^{3}-Leu^{4}-Lys^{5}-GLn^{6}-Ser^{7} V_H aminoterminal anchor region. The fragments overlapping the anchor regions were mutated to the original anchor region amino acids. The structure was then subjected to 100 steps of a steepest descent energy minimalization procedure (Fletcher and Reeves, 1964), fixing all atoms except those of the linker peptide. From the results of this procedure, it was concluded that the Glu-Ser dipeptide could be replaced by the original Lys^{107}-Arg^{108} V_L carboxyterminal amino

acids and that the Tyr-Val dipeptide could be replaced by the original Gln^1-Val^2 V_H aminoterminal amino acids. Thus, a linker peptide with the sequence: Ala-Gly-Gln-Gly-Ser-Cys-Val was derived. In order to prevent unwanted disulfide bridge formation, the peptide linker residue Cys was mutated to Ser, yielding the L12 peptide linker with sequence: Ala-Gly-Gln-Gly-Ser-Ser-Val.

The cDNA encoding the synthetic peptide linker L12 was then inserted between the cDNA encoding the V_L domain and the cDNA encoding the V_H domain of MA-15C5, resulting in the construction of the synthetic cDNA pMC5-K12A' as described below.

The 419 bp SmaI-HindIII fragment from pUC19-G_6 (Vandamme et al., 1990) was ligated in the EcoRI-HindIII treated pMa/c vector, in which the EcoRI recessing end was filled in with Klenow enzyme, yielding pMA/c-G6. A "TGAATTC" sequence was inserted in pMa/c-G_6 by site-directed mutagenesis between nucleotides 350 and 351 (on the pUC19-G_6 fragment sequence), introducing a TGA STOP codon at the presumed end of the J region of V_H and an additional EcoRI site. The resulting plasmid, pMa/c-G_{60}, was digested with EcoRI, treated with Klenow enzyme and religated, yielding pMa/c-G_0, in which the EcoRI restriction sites, together with the intervening sequences, are removed.

The 821 bp EcoRI-Bg1II fragment from pCMβDHFR-13-15C5kMu (Vandamme et al, 1990), containing the total kappa chain coding sequence (including the secretion signal) and 3' untranslated sequence, was inserted in EcoRI-BamHI digested pMA/c, yielding pMA/c-Kb. The 406 bp AccI-XbaI restriction fragment from pMA/c-G_0, of which the AccI recessing end was made blunt with Klenow enzyme, was transferred to StyI

(filled in)-XbaI treated pMa/c-Kb to yield pMa/c-KG_0. In this step, the 226 bp fragment comprising the carboxyterminal part of the kappa constant region (C_L) and the kappa 3' untranslated sequence was deleted. A single site-directed mutagenesis with the 72-mer oligodeoxynucleotide dCAAAGTTGGAAATCAAGCGTGCTGGTCAAGG-CTCTTCTGTTCAAGTTCAGCTGAAGCAGTCAGGACCTGGCC was performed on pMa/c-KG_0 to: i) delete the 328 bp DNA sequence separating the Arg^{108} of the kappa chain from the codon for Lys^5 of the gamma chain; ii) reintroduce cDNA sequence coding for amino acids 1 to 4 missing at the NH_2-terminus of V_H; and iii) insert the peptide linker L12 between the carboxyterminal end of V_L (Arg^{108}) and the aminoterminal end of V_H (Gln^1), yielding pMa/c-$K_{12}G_0$. 3 silent mutations were simultaneously introduced in the kappa coding sequence: the ATA Ile^{106} codon was changed to ATC; the AAA Lys^{107} codon was changed to AAG; and the CGG Arg^{108} codon was changed to CGT. The 818 bp EcoRI (filled in) - XbaI restriction fragment from pMa/c-$K_{12}G_0$ was then transferred to BamHI (filled in)-XbaI treated pVL1393 (British Biotechnology Ltd., Oxford, UK), yielding pVL-K12A'.

Sf9 cells were grown at 27°C in Grace's insect cell culture medium supplemented with 10% (vol/vol) fetal calf serum, 3.3% (vol/vol) yeastolate, and 3.3% (vol/vol) lactalbumin hydrolysate (TMNF medium) essentially as described by Summers and Smith (1987). The Sf9 cells (2 x 10^6 cells in a 25 cm^2 flask) were transfected with 1 µg AcNPV DNA and 10 µg pVL-K12A' by the Ca-phosphate co-precipitation method (Gorman et al., 1985), and the resulting culture supernatant was harvested 5-7 days later for cloning of recombinant baculovirus and for measurement of human fibrin fragment D-dimer binding protein in solid-phase enzyme-linked immunosorbent assay (ELISA).

For the cloning of recombinant baculovirus, fresh monolayers of Sf9 cells (1.5 x 10^6 Sf9 cells in a 6-well culture plate) were infected with eight 10-fold serial dilutions (between 10^3 and 10^{10}) of the cotransfection culture supernatant and subsequently overlaid with 1.5 percent low melting agarose containing 2-fold concentrated Grace's medium. When plaques were well formed (5-10 days post-infection), the putative recombinant plaques (occlusion-negative) were identified using a dissection microscope.

The recombinant plaques resuspended in 1 ml of TMNF medium and 50 µl aliquots were used to infect fresh monolayers of Sf9 cells (2 x 10^6 cells in a 25 cm^2 culture flask) overlaid with 4 ml TMNF medium. The resulting culture supernatants were harvested 48 h later for assessment of human fibrin fragment D-dimer binding in ELISA. The recombinant virus, AcNpVLK$_{12}$G$_0$, was then purified by 4 rounds of plaque purification. For each round, the concentration of fibrin fragment D-dimer binding protein was assessed in ELISA.

The purity of the isolated recombinant virus was confirmed in filter-hybridization experiments (Kafatos et al., 1979). In doing so, the DNA was extracted from 2 x 10^6 Sf9 insect cells and transferred to nitrocellulose. Hybridization was performed with either: a 36 bp probe (GACCCAGTCTCCATCTTCCATGTATGCATCT-CTAGG) complementary to the 72-107 bp aminoterminal sequence of the cDNA encoding the kappa chain of MA-15C5; or a probe complementary to the 134-167 bp polyhedrin cDNA sequence(CTACCCTCGACCCCCAAGACAACTACCTA-GTGGC) that is deleted in the pVL-K12A' cDNA by recombination.

For the large scale production of the SCA encoded by pMC5-K12A' and pVL-K12A', called "scFV-K$_{12}$G$_0$", 40 x 10^6 Sf9 cells in 175 cm^2 culture flasks were infected

with 200 x 10^6 plaque forming units of recombinant virus AcNpVLK$_{12}$G$_0$. After incubation for 48 h at 27°C, the conditioned medium, containing up to 15 µg scFv-K$_{12}$G$_0$ per ml, but on average approximately 4.5 µg/ml, was removed and centrifuged at 1,000xg for removal of cell debris.

scFv-K$_{12}$G$_0$ was purified by ion exchange chromatography on SP-Sephadex followed by gel filtration on Sephadex-G100 superfine. In the first step, 1.5 l of conditioned medium with a pH adjusted to 4.5 was applied on a 0.9 x 2 cm SP-Sephadex column equilibrated with 0.05 M NaH_2PO_4, pH 4.5, containing 0.05 M NaCl, 0.01% Tween 80 and 10 KIU/ml aprotinin. Elution was performed with a 60 ml gradient from 0.05 M to 1.0 M NaCl in 0.05 M NaH_2PO_4, pH 5.5. The fractions containing scFv-K$_{12}$G$_0$, as measured in ELISA specific for fibrin fragment D-dimer binding protein, were pooled and the pH was increased to 7.4 with 1M NaOH. The pooled fractions (representing 7 ml with a concentration of 0.7 mg scFv-K$_{12}$G$_0$ per ml) were concentrated on a Centricon 10 microconcentrator (Amicon) to a final volume of 0.05 ml. The concentrated sample was applied on a 1.0 x 110 cm Sephadex-G100 superfine column equilibrated with 0.02 M Tris-HCl buffer, pH 7.4, containing 0.3 M NaCl, 0.01% Tween 80 and 10 KIU/ml aprotinin. The fractions containing scFv-K$_{12}$G$_0$ were pooled and found to migrate to a single 25,500 Mr band on reduced SDS gel electrophoresis.

NH_2-terminal amino acid analysis of the scFv-K$_{12}$G$_0$, so obtained, revealed that the MA-15C5 kappa signal peptide was cleaved off by the insect cells just in front of mature kappa Asp^1 residue. The scFv-K$_{12}$G$_0$ was also found to bind to immobilized D-dimer with an affinity constant of 4 x 10^9 M^{-1}, as compared to 2.0 x 10^{10} m^{-1} for intact MA-15C5. This finding indicates

that, in scFv-$K_{12}G_0$, the MA-15C5 V_L and V_H domains can reassociate efficiently, resulting in the reconstitution of an intact, functionally active, antigen binding site. Further, it can be concluded that, provided the first linker peptide which connects the V_L and V_H domains does not put any spatial distance or structural constraints on the overall β-sheet structure of the framework regions, the molecular interactions responsible for the conserved framework structure will assure the proper folding of the hypervariable domain loops in the antigen binding site in the Fv fragment.

When injected as a bolus (2.8 μg/kg), scFv-$K_{12}G_0$ was cleared from the plasma of rabbits with a half-life of 10 minutes and a clearance rate of 5.1 ml/min^{-1}, as compared to 90 minutes and 210 ml/min^{-1} for intact MA-15C5. These results indicate that scFv-$K_{12}G_0$ can be useful for targeting radioisotopes or plasminogen activators to blood clots in vivo.

Example 9 : Synthesis and Expression of a pVL-K12A-PA-II' as in Example 3

A transfer vector pVL-K12A-PA-II', encoding the SCAPA called "$K_{12}G_0S_{32}$", for expression in Sf9 insect cells was constructed starting from the plasmids pMA/c-$K_{12}G_0$ and pVL-K12A' of Example 8, pULscu-PA (Nelles et al., 1987) and the pMA/c mutagenesis vector. A 1117 bp NcoI-HindIII fragment from pULscu-PA, containing the sequence encoding a scuPA fragment consisting of amino acids 67 to 411, connected to the 3'-untranslated sequence, was ligated in the BamHI-HindIII treated pMa/c mutagenesis vector, yielding pMa/c-scu-PA'.

By site-directed mutagenesis, the following mutations were introduced in the DNA sequence of scuPA

(nucleotide numbering refers to Fig. 5; amino acid numbering refers to the Holmes et al. (1985) scuPA sequence) to yield pMa/C-scu-PA'm. The Lys^{135}-Lys^{136} plasmid cleavage site in urokinase was removed by substituting Lys^{135} with Gln, using the 17-mer oligonucleotide dGGGCTTTTGTCCATCTG (underlined residue differs from the wild type residue). The Arg^{156}-Phe^{157} thrombin cleavage site in urokinase was removed by mutating Phe^{157} to Asp (nucleotides 628-630) with the 31-mer oligonucleotide dCCCAATAATCTTATCGCGAGGCCTCAGAG-TC. To facilitate the screening of the mutants, a StuI restriction site (nucleotides 619-624) was simultaneously created by changing the CCC Pro^{155} codon to CCT. The 33-mer oligonucleotide dGACAGAGCCCCCGCGGTGACGACGGTAGATGGC was used to modify 3 Arg codons: Arg178 and Arg^{179} AGG rare codons (nucleotides 691-696) were replaced by CGT codons, while for screening purposes, the Arg181 CGG codon was changed to CGC, generating a SacII restriction site (nucleotides 699-704). The BamHI restriction site in urokinase (nucleotides 1352-1357) was deleted by changing the ATC Ile^{399} codon to ATT with the 18-mer oligonucleotide dGTGACTGCGAATCCAGGG. This mutation was performed to facilitate further manipulation of the chimeric cDNA, using the BamHI restriction site present in the variable kappa light-chain coding sequence. One of the two FspI restriction sites (nucleotides 963-968) was removed by changing the GCG Ala^{369} codon to GCA with the 20-mer oligonucleotide dCGGGATGGCTGTGCACACCT. The FspI enzyme cleaves the remaining site precisely in front of amino acid Ala^{132}, which was used as the NH_2-terminal amino acid of the truncated scuPA. To facilitate further manipulation of the chimeric gene, one of the two EcoRI restriction sites (nucleotides 646-651) was deleted by changing the GAA Glu^{163} codon to

GAG with the 30-mer oligonucleotide dGATGGTGGTGAACTCTCCCCCAATAATCTT, and the PvuII restriction site (nucleotides 1090-1095) was removed by changing the CAG Gln^{311} codon to CAA with the 19-mer oligonucleotide dGTCATTTTCAGTTGCTCCG.

The 613 bp BamHI-HindIII fragment from pMa/c-$K_{12}G_0$, which encodes the carboxyterminal sequence of scFv-$K_{12}G_0$ and of which the HindIII end was filled in with Klenow enzyme, was ligated in BamHI-FspI treated pMa/c-scu-PA'm. The resulting plasmid pMa/c-12VS contained the sequence encoding the carboxyterminal region of scFv-$K_{12}G_0$ in front, but out of frame, of the aminoterminal sequence of the truncated catalytic domain of scuPA. Deletion oligonucleotide-directed mutagenesis was performed on pMa/c-12VS to delete the 22 nucleotides that still separated the carboxyterminal amino acid (Ser^{232}) of scFv-$K_{12}G_0$ and the first amino acid (Ala^{132}) of the truncated catalytic domain of scuPA, yielding pMa/c-$12G_0S_{32}$. The 51-mer oligonucleotide dAGAGGAGGGCTTTTGTCCATCTGCTGAGGAGACGGTGACTGAGGTTCCTTG used was complementary to the 9 carboxyterminal amino acids of the scFv-$K_{12}G_0$ molecule and to the 8 aminoterminal amino acids of the low molecular weight form (truncated catalytic domain) of scuPA. Finally, to reconstruct the total scFv-$K_{12}G_0$, domain of the chimeric molecule, the BamHI-XbaI fragment of pVL$K_{12}G_0$, containing the carboxyterminal $K_{12}G_0$ coding sequence, was replaced by the 1521 bp BamHI-XbaI restriction fragment from pMA/c-$12G_0S_{32}$, yielding pVL-K12A-PA-II'.

Sf9 cells were grown at 27°C in Grace's insect cell culture medium supplemented with 10% (vol/vol) fetal calf serum, 3.3% (vol/vol) yeastolate, and 3.3% (vol/vol) lactalbumin hydrolysate (TMNF medium), essentially as described by Summers and Smith (1987). The Sf9 cells (2 x 10^6 cells in a 25 cm^2 flask) were

transfected with 1 μg AcNPV DNA and 10 μg pVL-K12A-PA-II' by the Ca-phosphate co-precipitation method (Gorman et al, 1985), and the resulting culture supernatant was harvested 5-7 days later for cloning of recombinant baculovirus and for assessment of human fibrin fragment D-dimer binding protein in solid-phase enzyme-linked immunoassay (ELISA).

For the cloning of recombinant baculovirus, fresh monolayers of Sf9 cells (1.5 x 10^6 Sf9 cells in a 6-well culture plate) were infected with eight 10-fold serial dilutions (between 10^3 and 10^{10}) of the cotransfection culture supernatant and subsequently overlaid with 1.5% low melting agarose containing 2-fold concentrated Grace's medium. When plaques were well formed (5-10 days post-infection), the putative recombinant plaques (occlusive-negative) were identified using a dissection microscope (Summers and Smith, 1987).

The recombinant plaques were resuspended in 1 ml of TMNF medium, and 50 ul aliquots were used to infect fresh monolayers of Sf9 cells (2 x 10^6 cells in a 25 cm^2 culture flask) overlaid with 4 ml TMNF medium. The resulting culture supernatants were harvested 48 h later for assessment of human fibrin fragment D-dimer binding protein in ELISA.

The recombinant virus ($AcNpVLK_{12}G_0S_{32}$) was then purified by 4 rounds of plaque purification. For each round, the expression of fragment D-dimer binding protein and of uPA-related antigen was assessed in ELISA. The purity of the isolated recombinant virus was confirmed in filter-hybridization experiments (Kafatos et al., 1979).

For the large scale production of $K_{12}G_0S_{32}$, 40 x 10^6 Sf9 cells in 175 cm^2 culture flasks were infected with 200 x 10^6 plaque forming units of recombinant virus

AcNpVLK$_{12}$G$_0$S$_{32}$. After incubation for 48 h at 27°C, the conditioned medium was removed and centrifuged at 1,000 g for removal of cell debris.

K$_{12}$G$_0$S$_{32}$ was purified as described in Example 6 by ion exchange chromatography on SP-Sephadex followed by gel filtration on Sephadex-G100 superfine.

The specific activity of the resulting K$_{12}$G$_0$S$_{32}$ towards a chromogenic substrate for urokinase was $\leq$ 1,000 IU/mg before and 100,000 IU/mg uPA equivalent after conversion to its two-chain derivative with plasmin. The specific activity of both the single-chain and two-chain form on fibrin plates was 100,000 IU/mg uPA equivalent. Activation of plaminogen by K$_{12}$G$_0$S$_{32}$ obeyed Michaelis-Menten kinetics with Km= 2.9 $\pm$ 0.6 μM and a k_2= 3.7 $\pm$ 0.6 s^{-1} (mean $\pm$ SD; n= 3), as compared to Km= 12 μM and k_2= 4.8 s^{-1} for recombinant scuPA-32k (low M_r scuPA consisting of amino acids Leu^{144} to Leu^{411}).

Single-chain K$_{12}$G$_0$S$_{32}$ induced a dose- and time-dependent lysis of a ^{125}I-fibrin labeled human plasma clot immersed in citrated human plasma; fifty percent lysis in 2 h was obtained with 0.70 $\pm$ 0.07 μg/ml (mean $\pm$ SD; n= 5) as compared to 8.8 $\pm$ 0.1 μg/ml for recombinant scuPA-32k (mean $\pm$ SD; n= 3).

With two-chain K$_{12}$G$_0$S$_{32}$, fifty percent clot lysis in 2 h required 0.25 $\pm$ 0.03 μg/ml (mean $\pm$ SD; n= 3) as compared to only 0.62 $\pm$ 0.04 μg/ml (mean $\pm$ SD; n= 2) for recombinant tcuPA-32k. Fragment D-dimer inhibited the fibrinolytic activity of K$_{12}$G$_0$S$_{32}$ (50 percent inhibition with 6 μg fragment D-dimer/ml) but not of scuPA-32k.

These results indicate that low M_r scuPA of this invention can be targeted to a fibrin clot with a single-chain Fv fragment of a fibrin-specific antibody, resulting in a 13-fold increase of the fibrinolytic

potency of the single-chain form and a 2.5-fold increase of the potency of the two-chain form, as compared to that of their uPA-32k counterparts.

Needless to say, this invention is not limited to the transformation of a specific host microorganism or the use, for this purpose, of a chimaeric gene containing any specific promoter, signal sequence, sca or scapa gene and/or 3' transcription regulation sequence of this invention, or the use of any specific SCA or SCAPA, expressed by such a transformed host, for the specific purposes mentioned above. In this regard, equivalents of the foregoing Examples will be readily apparent to those skilled in the art in view of the disclosure herein of the invention. For example, the DNA sequences of the described sca and scapa genes (and consequently the amino acid sequences of the resulting SCAs and SCAPAs) can be easily modified by: 1) replacing some codons with others that code either for the same amino acids or for other amino acids; and/or 2) deleting or adding some codons; provided that such modifications do not substantially alter the biological properties of the encoded SCAs or SCAPAs.

Also this invention is not limited to an SCA or an SCAPA derived from a monoclonal antibody directed to fibrin or fibrin D-links, such as MA-15C5 antibody. This invention encompasses SCAs and SCAPAS derived from monoclonal antibodies directed to other thrombus constituents such as: a) antibodies to blood platelets, for example antibodies to resting and activated platelet surface receptors, e.g., antibodies to platelet membrane glycoprotein IIb/IIIa (Bode et al., 1990) or antibodies (e.g., MA-libs-1) specific for ligand-occupied receptor conformers (Frelinger et al., 1990); or b) antibodies to alpha 2-antiplasmin. This invention also encompasses SCAs and SCAPAs derived from

other monoclonal antibodies directed to fibrin such as the 59D8 antibodies (Bode et al., 1987).

References

- Albers et al. (1989) Molecular Biology of the Cell, Garland Publishing Inc, New York, London
- Artelt et al. (1988) Gene 69:213-219
- Astrup et al. (1952) Arch. Biochem. Biophys. 40:346
- Bachmann (1987) In "Haemostasis and Thrombosis: Basic Principles and Clinical Practice", Colman R.W., Hirsch J, Marder V.J. and Salzman E.W. (eds.), J.B. Lippincott Company, pp. 318-339
- Baker et al. (1988) Gene 69:349-355
- Beaucage and Caruthers (1981) Tetrahedron Letters 22:1859-1862
- Bergman et al. (1983) Science 220:1181
- Bernard et al. (1979) Gene 5:59-79
- Bernstein et al. (1977) J.Mol.Biol. 112:535-542
- Betzel et al. (1988) Acta Crystallographica B 44:163
- Bird et al. (1988) Science 242:423-426
- Blundell et al. (1987) Nature 326:347-352
- Bode et al. (1987) J. Mol. Cell. Cardiol. 19:335-341
- Bode et al. (1990) Circulation, supp. III, 82:376
- Chaudhary et al. (1989) Nature 339:394-397

- Chaudhary et al. (1990) Proc. Natl. Acad. Sci. USA 87:1066-1070

- Claessens et al. (1989) Protein Engineering 2:335-345

- Collen et al. (1983) J. Clin. Invest. 71:368

- Collen et al. (1989) Fibrinolysis 3:197-202

- Collen et al. (1990) Cir. 82:1744-1753

- Colman et al. (1987) In "Haemostasis and Thrombosis: Basic Principles and Clinical Practice", Colman R.W., Hirsch J, Marder V.J. and Salzman E.W. (eds.), J.B. Lippincott Company, pp. 3-17

- Darras et al. (1986) Thrombosis and Haemostasis 56:411

- De Boer et al. (1983) Proc. Natl. Acad. Sci. USA 80:21-25

- Delhaise et al. (1985) J. Mol. Graph. 3:116-119

- Dewerchin et al. (1990) Eur. J. Biochem. 185:141-149

- Elms et al. (1983) Thrombosis and Haemostasis 50:591-594

- Flameng et al. (1985) J.Clin.Invest. 75:84

- Fletcher and Reeves (1964) Comput.J. 6:163-168

- Freiman (1987) In "Haemostasis and Thrombosis: Basic Principles and Clinical Practice", Colman

R.W., Hirsch J, Marder V.J. and Salzman E.W. (eds.), J.B. Lippincott Company, pp. 1123-1135

- Frelinger et al. (1990) J. Biol. Chem. 265:6346 et seq.

- Fuller (1982) Gene 19:43-54

- Gorman (1985) In: DNA Cloning. A Practical Approach, Vol. II, Glover, D.M. (ed.), IRL Press, Oxford-Washington, pp. 143-190

- Gurewich et al. (1984) J. Clin. Invest. 73:1731-1739

- Haber et al. (1989) Science 243:51-56

- Haber (1990) Cir. 82:1874-1876

- Hantgan et al. (1987) In "Haemostasis and Thrombosis: Basic Principles and Clinical Practice", Colman R.W., Hirsch J, Marder V.J. and Salzman E.W. (eds.), J.B. Lippincott Company, pp. 269-288

- Harlow and Lane (1988) Antibodies, a laboratory manual, Cold Spring Harbor Laboratory

- Hogg et al. (1987) Mol. Immunol. 24:797

- Holmes et al. (1985) Bio/Technology 3:923-929

- Holvoet et al. (1989) Thrombosis and Haemostasis 61:307-313

- Hui et al. (1986) Hybridoma 3:215-222

- Huston et al. (1988) Proc. Natl. Acad. Sci. USA 85:5879-5883

- Jibrat et al. (1987) J. Mol. Biol. 198:425-443
- Kabat et al. (1987) "Sequences of Proteins of Immunological Interest", US department Health Human Services, Washington, DC
- Kafatos et al (1979) Nuc. Acids Res. 7:1541-1552
- Kramer and Fritz (1988) Meth. Enzymol. 154:350-367
- Kramer et al. (1984) Nucl. Acids Res. 12:9441-9456
- Kudryk et al. (1984) Mol. Immunol. 21:89-94
- Lee et al. (1981) Nature 294:228-232
- Lijnen et al. (1986) J. Biol. Chem. 261:1253
- Lijnen et al. (1988) J. Biol. Chem. 263:5594-5598
- LoBuglio et al. (1989) Proc. Natl. Acad. Sci. USA 86:4220-4224
- Luckow and Summers (1987) Bio/Technology 6:47-55
- Luckow and Summers (1989) Virology 170:31-39
- MacLachlan (1979) J. Mol. Biol. 128:49-79
- Madison et al. (1989) Nature 339:721-724
- Madison et al. (1990) Proc. Natl. Acad. Sci. USA 87:3530-3533
- Maeda (1989) Ann. Rev. Entomol. 34:351-372
- Marder and Bell (1987) In "Haemostasis and Thrombosis: basic principles and clinical practice", Colman R.W., Hirsch J, Marder V.J. and

Salzman E.W. (eds.), J.B. Lippincott Company, pp. 1393-1437

- Martial et al. (1979) Science 205:602-607
- Menck et al. (1987) Gene 53:21-29
- Michaelis et al. (1983) J.Bacteriol. 154:366
- Nelles et al. (1987) J. Biol. Chem. 262:5682-5689
- Norrander et al. (1983) Gene 26:101-106
- Oswald et al. (1989) Nature 337:579-582
- Pennica et al. (1983) Nature 301:214-221
- Riechmann et al. (1988) Nature 332:323-327
- Sambrook et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory
- Scheefers-Borchel et al. (1985) Proc. Natl. Acad. Sci. 82:7091-7095
- Sheriff et al. (1987) Proc. Natl. Acad. Sci. USA 84:8075-8079
- Sherry (1987) Am. J. Cardiol. 59:984-989
- Stanssens et al. (1987) Manual EMBO Laboratory Course 'Directed Mutagenesis and Protein Engineering" held at Max-Planck Institute für Biochemie, Martinsried, West-Germany, July 4-18, 1987
- Stanssens et al. (1989) Nucleic Acids Research 17:4441-4454

- Summers and Smith (1987) A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures. Bulletin 1555 Texas Agricultural Experiment Station.

- Tainer et al. (1982) J. Mol. Biol. 160:181

- Taylor (1988) Protein Engineering 2:77-86

- Verhoeyen et al. (1988) Science 239:1534-1536

- Vandamme et al. (1990) Eur. J. Biochem. 192:767-775

- Voller et al. (1976) In "Manual of Clinical Immunology" Eds. M.R. Rose and H. Friedman, American Society of Microbiology, Washington, pp. 506-512

- Von Heyne (1988) Biochem. Biophys. Acta 947-307-333

- Watson (1984) Nucleic Acids Research 12:5145-5164

- Wüthrich (1986) "NMR of Proteins and Nucleic Acids", J.Wiley and Sons

- Wyckoff et al., eds. (1985) in "Diffraction Methods for Biological Macromolecules", Meth. Enzymol. vols. 114 and 115. Academic Press.

- Yanisch-Perron et al. (1985) Gene 33:103-119

- Yasuda et al. (1989) J.Am.Coll.Cardiol. 13:1409

- Zabeau and Stanley (1982) EMBO J. 1:1217-1224

CLAIMS

1. A single-chain antibody: which is specific to a thrombus constituent, preferably fibrin; which preferably comprises all or especially a part of a monoclonal antibody directed against the thrombus constituent, especially fibrin, particularly fibrin cross-links, quite particularly fibrin D-dimer; which is preferably properly folded for binding to the thrombus constituent; and/or which is preferably glycosylated; the part of the monoclonal antibody preferably comprising all or especially a part of variable domains of the light and heavy chains of the monoclonal antibody, connected by means of a first linker peptide.

2. The single chain antibody of claim 1 which is derived from monoclonal antibody MA-15C5 and which preferably has one of the amino acid sequences shown in Fig. 6.

3. A thrombolytic agent, comprising the single-chain antibody of claim 1 or 2, connected, preferably by means of a second linker peptide, to a plasminogen activating portion, preferably a catalytic domain of urokinase or tissue plasminogen activator, preferably of scuPA.

4. The thrombolytic agent of claim 3, which has increased half-life and which: a) is in a non-glycosylated form or in a super-glycosylated form or in a form in which some glycosylation is added to, and other glycosylation is removed from, the thrombolytic agent; and/or b) is in a form which is resistant to a plasminogen activator inhibitor; and/or c) has all or at least a significant part of the A domain of its

plasminogen activating portion serving as the second linker peptide.

5. The thrombolytic agent of claim 3 or 4 in which: a) the plasminogen activating portion is the catalytic domain of urokinase, preferably having the amino acid sequence shown in Fig. 5 from Leu144 to Leu411; and b) in which the C-terminal end of the single-chain antibody is either directly connected to the N-terminal end of the catalytic domain or preferably is connected by the second linker peptide, derived from the urokinase region connecting the kringle and catalytic domains, preferably having the amino acid sequence shown in Fig. 5 from Ala132 to Glu143.

6. A DNA fragment coding for the single-chain antibody of claim 1 or 2 or the thrombolytic agent of anyone of claims 3-5.

7. A chimaeric gene for transforming a host cell to express the DNA fragment of claim 6, comprising the following operably linked DNA fragments in the same transcriptional unit: i) a promoter capable of directing the expression of the DNA fragment in the host cell, preferably an insect cell, a mammalian cell, or Escherichia coli cell, particularly an insect cell or a mammalian cell; ii) the DNA fragment of claim 6; iii) a suitable 3' transcription regulation sequence for the host cell; and optionally, between fragments i) and ii), iv) a signal sequence coding for a signal peptide capable of directing secretion of the expression product of the DNA fragment ii) from the host cell.

8. A host cell, preferably an insect cell, a mammalian cell, or Escherichia coli cell, particularly

an insect or mammalian cell, transformed with the chimaeric gene of claim 7.

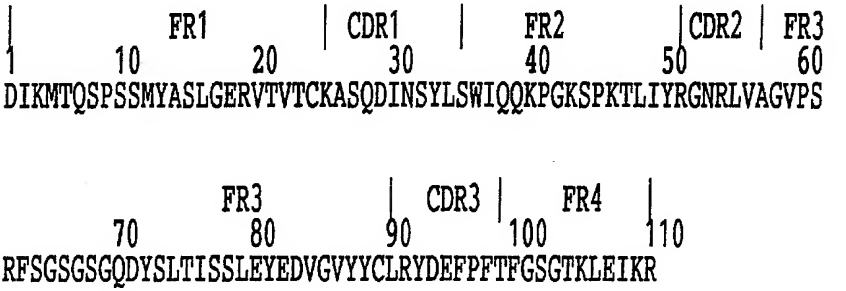
9. A method for producing the single-chain antibody of claim 1 or 2 or the thrombolytic agent of anyone of claims 3-5, comprising: culturing the host cells of claim 8; and recovering the antibody or thrombolytic agent from the culture medium.

10. A process, comprising the use of the single-chain antibody of claim 1 or 2 for imaging of a thrombus in a living organism.

11. A process, comprising the use of the thrombolytic agent of anyone of claims 3-5 for dissolving a thrombus in a living organism.

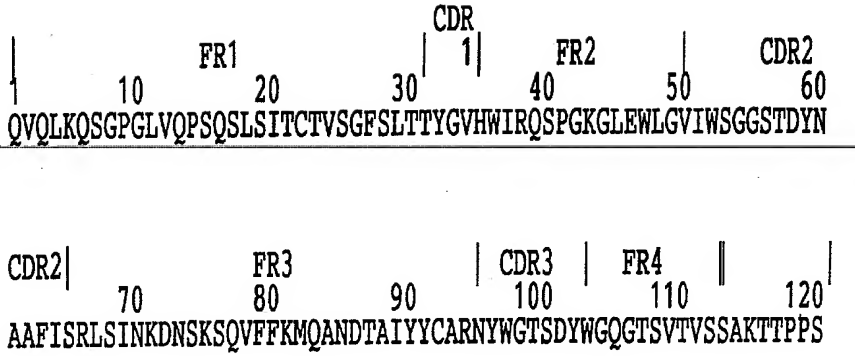
12. A vector for transforming the cell of claim 8, comprising the chimaeric gene of claim 7.

FIGURE 1



No insertions

FIGURE 2



Insertions

in FR3 : 82ABC = 82NSL

in CDR3: 100ABCDEFGHIJK = 100----------M

FIGURE 3

```
        10          20          30          40          50
EcoRI
GAATTCCGCA CATGAGGACC CCTGCTCAGT TTCTTGGAAT CTTGTTGCTC

        60          70  >       80          89
TGGTTTCCAG GTATCAAATG T GAC ATC AAG ATG ACC CAG TCT
                        Asp Ile Lys MET Thr Gln Ser

     98         107         116         125
CCA TCT TCC ATG TAT GCA TCT CTA GGA GAG AGA GTC ACT
Pro Ser Ser MET Tyr Ala Ser Leu Gly Glu Arg Val Thr

134         143         152         161         170
GTC ACT TGC AAG GCG AGT CAG GAC ATT AAT AGC TAT TTA
Val Thr Cys Lys Ala Ser Gln Asp Ile Asn Ser Tyr Leu

        179         188         197         206
      BamHI
AGC TGG ATC CAG CAG AAA CCA GGG AAA TCT CCT AAG ACC
Ser Trp Ile Gln Gln Lys Pro Gly Lys Ser Pro Lys Thr

    215         224         233         242
CTG ATC TAC CGT GGA AAC AGA TTG GTT GCT GGG GTC CCA
Leu Ile Tyr Arg Gly Asn Arg Leu Val Ala Gly Val Pro

251         260         269         278         287
TCA AGG TTC AGT GGC AGT GGA TCT GGG CAA GAT TAT TCT
Ser Arg Phe Ser Gly Ser Gly Ser Gly Gln Asp Tyr Ser
```

FIGURE 3 (continued 1)

```
        296         305         314         323
CTC ACC ATC AGC AGC CTG GAG TAT GAA GAT GTG GGA GTT
Leu Thr Ile Ser Ser Leu Glu Tyr Glu Asp Val Gly Val

    332         341         350         359
TAT TAT TGT CTA CGG TAT GAT GAG TTT CCA TTC ACG TTC
Tyr Tyr Cys Leu Arg Tyr Asp Glu Phe Pro Phe Thr Phe

368         377         386         395         405
GGC TCG GGG ACA AAG TTG GAA ATA AAA CGG GCTGATGCTG
Gly Ser Gly Thr Lys Leu Glu Ile Lys Arg

      415         425         435         445         455
CACCAACTGT ATCCATCTTC CCACCATCCA GTGAGCAGTT AACATCTGGA

      465         475         485         495         505
GGTGCCTCAG TCGTGTGCTT CTTGAACAAC TTCTACCCCA AAGACATCAA

      515         525         535         545         555
TGTCAAGTGG AAGATTGATG GCAGTGAACG ACAAAATGGC GTCCTGAACA

      565         575         585         595         605
GTTGGACTGA TCAGGACAGC AAAGACAGCA CCTACAGCAT GAGCAGCACC

      615         625         635         645         655
CTCACGTTGA CCAAGGACGA GTATGAACGA CATAACAGCT ATACCTGTGA

      665         675         685         695         705
GGCCACTCAC AAGACATCAA CTTCACCCAT TGTCAAGAGC TTCAACAGGA
```

FIGURE 3 (continued 2)

```
      715        725        735        745        755
ATGAGTGTTA GAGACAAAGG TCGGGCGAGC TCGAATTAAT TCACTCCTCA

      765        775        785        795        805
GGTGCAGGCT GCCTATCAGA AGGTGGTGGC TGGTGTGGCC AATGCCCTGG

      815        825
                 BglII
CTCACAAATA CCACTGAGAT CT
```

FIGURE 4

```
        10          20          29          38
PstI        )
CTGCAGGAAT G AAG CAG TCA GGA CCT GGC CTA GTG CAG
           Lys Gln Ser Gly Pro Gly Leu Val Gln

        47          56          65          74
CCC TCA CAG AGC CTG TCC ATC ACC TGC ACA GTC TCT GGT
Pro Ser Gln Ser Leu Ser Ile Thr Cys Thr Val Ser Gly

     83          92         101         110
TTC TCA TTA ACT ACC TAT GGT GTA CAC TGG ATT CGC CAG
Phe Ser Leu Thr Thr Tyr Gly Val His Trp Ile Arg Gln

119         128         137         146         155
TCT CCA GGA AAG GGT CTG GAG TGG CTG GGA GTG ATA TGG
Ser Pro Gly Lys Gly Leu Glu Trp Leu Gly Val Ile Trp

        164         173         182         191
AGT GGT GGA AGC ACA GAC TAT AAT GCA GCT TTC ATA TCC
Ser Gly Gly Ser Thr Asp Tyr Asn Ala Ala Phe Ile Ser

    200         209         218         227
AGA CTG AGC ATC AAC AAG GAC AAT TCC AAG AGC CAA GTT
Arg Leu Ser Ile Asn Lys Asp Asn Ser Lys Ser Gln Val

236         245         254         263         272
TTC TTT AAA ATG AAC AGT CTG CAA GCT AAT GAC ACA GCC
Phe Phe Lys MET Asn Ser Leu Gln Ala Asn Asp Thr Ala
```

FIGURE 4 (continued 1)

```
        281         290         299         308
ATA TAT TAC TGT GCC AGA AAT TAT TGG GGA ACC TCT ATG
Ile Tyr Tyr Cys Ala Arg Asn Tyr Trp Gly Thr Ser MET

    317         326         335         344
GAC TAC TGG GGT CAA GGA ACC TCA GTC ACC GTC TCC TCA
Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser

353         362         371         380         390
                                                EcoRI
GCC AAA ACG ACA CCC CCA TCT GTCTATCCA CTGGAATTCG
Ala Lys Thr Thr Pro Pro Ser

      400
    HindIII
ATATCAAGCTT
```

FIGURE 5

```
        10          20          30          40          50
HindIII
AAGCTTCGGG CCAGGGTCCA CCTGTCCCCG CAGCGCCGTC GCGCCCTCCT

        60          70          80          90          99
GCCGCAGGCC ACCGAGGCCG CCGCCGTCTA GCGCCCCGAC CTCGCCACC

       108         117         126         135
ATG AGA GCC CTG CTG GCG CGC CTG CTT CTC TGC GTC CTG
MET Arg Ala Leu Leu Ala Arg Leu Leu Leu Cys Val Leu

    144         153         162         171
GTC GTG AGC GAC TCC AAA GGC AGC AAT GAA CTT CAT CAA
Val Val Ser Asp Ser Lys Gly Ser Asn Glu Leu His Gln

180         189         198         207         216
GTT CCA TCG AAC TGT GAC TGT CTA AAT GGA GGA ACA TGT
Val Pro Ser Asn Cys Asp Cys Leu Asn Gly Gly Thr Cys

       225         234         243         252
GTG TCC AAC AAG TAC TTC TCC AAC ATT CAC TGG TGC AAC
Val Ser Asn Lys Tyr Phe Ser Asn Ile His Trp Cys Asn

    261         270         279         288
TGC CCA AAG AAA TTC GGA GGG CAG CAC TGT GAA ATA GAT
Cys Pro Lys Lys Phe Gly Gly Gln His Cys Glu Ile Asp

297         306         315         324         333
AAG TCA AAA ACC TGC TAT GAG GGG AAT GGT CAC TTT TAC
Lys Ser Lys Thr Cys Tyr Glu Gly Asn Gly His Phe Tyr
```

FIGURE 5 (continued 1)

```
        342         351         360         369
                                NcoI
CGA GGA AAG GCC AGC ACT GAC ACC ATG GGC CGG CCC TGC
Arg Gly Lys Ala Ser Thr Asp Thr MET Gly Arg Pro Cys

    378         387         396         405
CTG CCC TGG AAC TCT GCC ACT GTC CTT CAG CAA ACG TAC
Leu Pro Trp Asn Ser Ala Thr Val Leu Gln Gln Thr Tyr

414         423         432         441         450
CAT GCC CAC AGA TCT GAT GCT CTT CAG CTG GGC CTG GGG
His Ala His Arg Ser Asp Ala Leu Gln Leu Gly Leu Gly

        459         468         477         486
AAA CAT AAT TAC TGC AGG AAC CCA GAC AAC CGG AGG CGA
Lys His Asn Tyr Cys Arg Asn Pro Asp Asn Arg Arg Arg

    495         504         513         522
CCC TGG TGC TAT GTG CAG GTG GGC CTA AAG CCG CTT GTC
Pro Trp Cys Tyr Val Gln Val Gly Leu Lys Pro Leu Val

531         540         549         558         567
                                FspI
CAA GAG TGC ATG GTG CAT GAC TGC GCA GAT GGA AAA|AAG
Gln Glu Cys MET Val His Asp Cys Ala Asp Gly Lys|Lys

        576         585         594         603
CCC TCC TCT CCT CCA GAA GAA|TTA AAA TTT CAG TGT GGC
Pro Ser Ser Pro Pro Glu Glu|Leu Lys Phe Gln Cys Gly
```

FIGURE 5 (continued 2)

```
    612         621         630         639
CAA AAG ACT CTG AGG CCC CGC TTT AAG ATT ATT GGG GGA
Gln Lys Thr Leu Arg Pro Arg Phe Lys Ile Ile Gly Gly

648         657         666         675         684
GAA TTC ACC ACC ATC GAG AAC CAG CCC TGG TTT GCG GCC
Glu Phe Thr Thr Ile Glu Asn Gln Pro Trp Phe Ala Ala

        693         702         711         720
ATC TAC AGG AGG CAC CGG GGG GGC TCT GTC ACC TAC GTG
Ile Tyr Arg Arg His Arg Gly Gly Ser Val Thr Tyr Val

    729         738         747         756
TGT GGA GGC AGC CTC ATC AGC CCT TGC TGG GTG ATC AGC
Cys Gly Gly Ser Leu Ile Ser Pro Cys Trp Val Ile Ser

765         774         783         792         801
GCC ACA CAC TGC TTC ATT GAT TAC CCA AAG AAG GAG GAC
Ala Thr His Cys Phe Ile Asp Tyr Pro Lys Lys Glu Asp

        810         819         828         837
TAC ATC GTC TAC CTG GGT CGC TCA AGG CTT AAC TCC AAC
Tyr Ile Val Tyr Leu Gly Arg Ser Arg Leu Asn Ser Asn

    846         855         864         873
ACG CAA GGG GAG ATG AAG TTT GAG GTG GAA AAC CTC ATC
Thr Gln Gly Glu MET Lys Phe Glu Val Glu Asn Leu Ile
```

FIGURE 5 (continued 3)

```
882         891         900         909         918
CTA CAC AAG GAC TAC AGC GCT GAC ACG CTT GCT CAC CAC
Leu His Lys Asp Tyr Ser Ala Asp Thr Leu Ala His His

        927         936         945         954
AAT GAC ATT GCC TTG CTG AAG ATC CGT TCC AAG GAG GGC
Asn Asp Ile Ala Leu Leu Lys Ile Arg Ser Lys Glu Gly

    963         972         981         990
      FspI
AGG TGT GCG CAG CCA TCC CGG ACT ATA CAG ACC ATC TGC
Arg Cys Ala Gln Pro Ser Arg Thr Ile Gln Thr Ile Cys

999        1008        1017        1026        1035
CTG CCC TCG ATG TAT AAC GAT CCC CAG TTT GGC ACA AGC
Leu Pro Ser MET Tyr Asn Asp Pro Gln Phe Gly Thr Ser

      1044        1053        1062        1071
TGT GAG ATC ACT GGC TTT GGA AAA GAG AAT TCT ACC GAC
Cys Glu Ile Thr Gly Phe Gly Lys Glu Asn Ser Thr Asp

   1080        1089        1098        1107
TAT CTC TAT CCG GAG CAG CTG AAA ATG ACT GTT GTG AAG
Tyr Leu Tyr Pro Glu Gln Leu Lys MET Thr Val Val Lys

1116        1125        1134        1143        1152
CTG ATT TCC CAC CGG GAG TGT CAG CAG CCC CAC TAC TAC
Leu Ile Ser His Arg Glu Cys Gln Gln Pro His Tyr Tyr
```

FIGURE 5 (continued 4)

```
      1161       1170       1179       1188
GGC TCT GAA GTC ACC ACC AAA ATG CTG TGT GCT GCT GAC
Gly Ser Glu Val Thr Thr Lys MET Leu Cys Ala Ala Asp

   1197       1206       1215       1224
CCA CAG TGG AAA ACA GAT TCC TGC CAG GGA GAC TCA GGG
Pro Gln Trp Lys Thr Asp Ser Cys Gln Gly Asp Ser Gly

1233       1242       1251       1260       1269
GGA CCC CTC GTC TGT TCC CTC CAA GGC CGC ATG ACT TTG
Gly Pro Leu Val Cys Ser Leu Gln Gly Arg MET Thr Leu

      1278       1287       1296       1305
ACT GGA ATT GTG AGC TGG GGC CGT GGA TGT GCC CTG AAG
Thr Gly Ile Val Ser Trp Gly Arg Gly Cys Ala Leu Lys

   1314       1323       1332       1341
GAC AAG CCA GGC GTC TAC ACG AGA GTC TCA CAC TTC TTA
Asp Lys Pro Gly Val Tyr Thr Arg Val Ser His Phe Leu

1350       1359       1368       1377       1386
      BamHI
CCC TGG ATC CGC AGT CAC ACC AAG GAA GAG AAT GGC CTG
Pro Trp Ile Arg Ser His Thr Lys Glu Glu Asn Gly Leu

      1395       1405       1415       1425       1435
GCC CTC TGA GGGTCCCCAG GGAGGAAACG GGCACCACCC GCTTTCTTGC
Ala Leu  .
```

FIGURE 5 (continued 5)

```
      1445       1455       1465       1475       1485
                                       HindIII
TGGTTGTCAT TTTTGCAGTA GAGTCATCTC CATCAGAAGC TTTTGGGGAG

      1495       1505       1515       1525       1535
CAGAGACACT AACGACTTCA GGGCAGGGCT CTGATATTCC ATGAATGTAT

      1545       1555       1565       1575       1585
CAGGAAATAT ATATGTGTGT GTATGTTTGC ACACTTGTTG TGTGGGCTGT

      1595       1605       1615       1625       1635
GAGTGTAAGT GTGAGTAAGA GCTGGTGTCT GATTGTTAAG TCTAAATATT

      1645       1655       1665       1675       1685
TCCTTAAACT GTGTGGACTG TGATGCCACA CAGAGTGGTC TTTCTGGAGA

      1695       1705       1715       1725       1735
GGTTATAGGT CACTCCTGGG GCCTCTTGGG TCCCCCACGT GACAGTGCCT

      1745       1755       1765       1775       1785
GGGAATGTAC TTATTCTGCA GCATGACCTG TGACCAGCAC TGTCTCAGTT

      1795       1805       1815       1825       1835
TCACTTTCAC ATAGATGTCC CTTTCTTGGC CAGTTATCCC TTCCTTTTAG

      1845       1855       1865       1875       1885
CCTAGTTCAT CCAATCCTCA CTGGGTGGGG TGAGGACCAC TCCTTACACT
```

FIGURE 5 (continued 6)

```
      1895       1905       1915       1925       1935
GAATATTTAT ATTTCACTAT TTTTATTTAT ATTTTTGTAA TTTTAAATAA

      1945       1955       1964
AAGTGATCAA TAAAATGTGA TTTTTCTGA
```

FIGURE 6

V_L-L_{ab}-V_H constructions

Construction 1.
[D1...G99-SGTKL-EIKR]-(AGQGSSV)-{QV-QLKQS-G8...S113}

Construction 2.
[D1...G99-SGTKL-EG*KG*]-(AGQGSSV)-{QV-QLKQS-G8...S113}

Construction 3.
[D1...G99-SGTKL-ES*KS*]-(AGQGSSV)-{QV-QLKQS-G8...S113}

Construction 4.
[D1...G99-SGTKL-ES*KS*]-(AGR(nG)GSSV)-{QV-QLKQS-G8...S113}
with n=0 to 4

Construction 5
[D1...K103-L]-(KESGSVSSEQLAQFRSLD)-{V-Q3...S113}

Construction 1A.
[D1...G99-SGTKL-EIKR]-(AGQGSSV)-{QV-QLKQS-G8...S120}

Construction 2A.
[D1...G99-SGTKL-EG*KG*]-(AGQGSSV)-{QV-QLKQS-G8...S120}

Construction 3A.
[D1...G99-SGTKL-ES*KS*]-(AGQGSSV)-{QV-QLKQS-G8...S120}

Construction 4A.
[D1...G99-SGTKL-ES*KS*]-(AGR(nG)GSSV)-{QV-QLKQS-G8...S120}
with n=0 to 4

Construction 5A
[D1...K103-L]-(KESGSVSSEQLAQFRSLD)-{V-Q3...S120}

FIGURE 6 (continued 1)

V_H-L_{ab}-V_L constructions

Construction 6.
{Q1...T107-SVTV-SS}-(DHGSHSHQSGGSGSG)-[DIK-M4...R108]

Construction 7.
{Q1...T107-SVTV-SS}-(DHGSHSEQSGGSGSG)-[DIK-M4...R108]

Construction 8.
{Q1...T107-SVTV-SS}-(GGGSHSEQSGGSGSG)-[DIK-M4...R108]

Construction 9.
{Q1...T107-SVTV-SS}-(GGGSGSGGSGGSGSG)-[DIK-M4...R108]

Construction 10.
{Q1...T107-SVTV-SS}-(GGGGSGGGGSGGGGS)-[DIK-M4...R108]

FIGURE 7

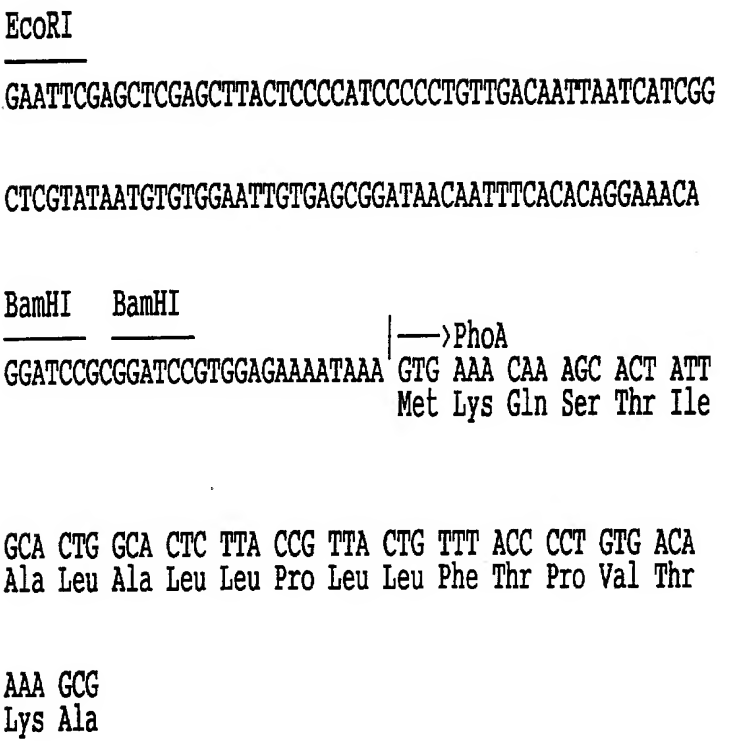

EcoRI
GAATTCGAGCTCGAGCTTACTCCCCATCCCCCTGTTGACAATTAATCATCGG
CTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACA
BamHI
BamHI
PhoA
GGATCCGCGGATCCGTGGAGAAAATAAA GTG AAA CAA AGC ACT ATT
Met Lys Gln Ser Thr Ile
GCA CTG GCA CTC TTA CCG TTA CTG TTT ACC CCT GTG ACA
Ala Leu Ala Leu Leu Pro Leu Leu Phe Thr Pro Val Thr
AAA GCG
Lys Ala

FIGURE 8

A.

ATG AGG ACC CCT GCT CAG TTT CTT GGA ATC TTG
Met Arg Thr Pro Ala Gln Phe Leu Gly Ile Leu

TTG CTC TGG TTT CCA GGT ATC AAA TGT
Leu Leu Trp Phe Pro Gly Ile Lys Cys

B.

ATG GCT GTC TTA GGG CTG CTC TTC TGC CTA GTG
Met Ala Val Leu Gly Leu Leu Phe Cys Leu Val

ACA TTC CCA AGC TGT GTC CTA TCC
Thr Phe Pro Ser Cys Val Leu Ser

FIGURE 9

A. $L_{ab}12$

```
            11          20          29          38
CA AAG TTG GAA ATA AAA CGG GCA GGC CAA GGG AGC TCA
   Lys Leu Glu Ile Lys Arg Ala Gly Gln Gly Ser Ser

        47          56          65
GTA CAA GTA CAA CTA AAG CAG TCA GGA CCT GGC C
Val Gln Val Gln Leu Lys Gln Ser Gly Pro Gly
```

B. $L_{ab}14$

```
            11          20          29          38
CG GGG ACA AAG TTG AAA GAA TCA GGA TCA GTC TCG AGT
   Gly Thr Lys Leu Lys Glu Ser Gly Ser Val Ser Ser

        47          56          65          74
GAA CAA TTA GCA CAA TTT AGA TCT TTA GAT GTA CAG CTG
Glu Gln Leu Ala Gln Phe Arg Ser Leu Asp Val Gln Leu

    83          92
AAG CAG TCA GGA CCT GG
Lys Gln Ser Gly Pro
```

FIGURE 9 (continued 1)

C. $L_{ab}15$

```
            11          20          29          38
CG GGG ACA AAG TTG GAA TCA AAG AGC GCT GGC CGC GGC
   Gly Thr Lys Leu Glu Ser Lys Ser Ala Gly Arg Gly

         47          56          65          74
GGC GGC GGC GGC TCG AGT GTC CAA GTA CAG CTG AAG CAG
Gly Gly Gly Gly Ser Ser Val Gln Val Gln Leu Lys Gln

     83
TCA GGA CCT GGC
Ser Gly Pro Gly
```

INTERNATIONAL SEARCH REPORT

International Application No PCT/EP 91/00767

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[5]: C 07 K 15/28, A 61 K 37/02, C 12 N 15/13, C 12 P 21/08

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[5]	C 07 K, C 12 N, C 12 P, A 61 K

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [10]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
X	EP, A, 0271227 (THE GENERAL HOSPITAL CORP.) 15 June 1988 see paragraph bridging pages 28 and 29	1,3,6-9,12
Y		2
	--	
Y	Thrombosis and Haemostasis, vol. 61, no. 2, 25 April 1989, P. Holvoet et al.: "Binding properties of monoclonal antibodies against human fragment D-dimer of cross-linked fibrin to human plasma clots in an in vivo model in rabbits", pages 307-313 see the whole article	2
	--	
A	EP, A, 0347078 (CELLTECH LTD) 20 December 1989 see page 4, lines 21-40	1-9,12
	--	
	./.	

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
8th August 1991	11 SEP 1991
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	[illegible] VAZELAAS

Form PCT/ISA/210 (second sheet) (January 1985)

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

A	Trends in Biotechnology, vol. 6, no. 2, February 1988, Elsevier Publications, (Cambridge, GB), G. Williams: "Novel antibody reagents: production and potential", pages 36-39,42 see page 39: "Single chain antibodies"	1-9,12

V. ☐ OBSERVATION WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This International search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim numbers 10,11 because they relate to subject matter not required to be searched by this Authority, namely:

   See PCT-Rule 39.1 (iv): methods for treatment of the human or animal body by surgery or therapy as well as diagnostic methods

2. ☐ Claim numbers because they relate to parts of the International application that do not comply with the prescribed requirements to such an extent that no meaningful International search can be carried out, specifically:

3. ☐ Claim numbers because they are dependent claims and are not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. ☐ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this International application as follows:

1 ☐ As all required additional search fees were timely paid by the applicant, this International search report covers all searchable claims of the International application

2 ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the International application for which fees were paid, specifically claims:

3 ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4 ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee

Remark on Protest

☐ The additional search fees were accompanied by applicant's protest

☐ No protest accompanied the payment of additional search fees

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. EP 9100767 SA 46812

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 06/09/91
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A- 0271227	15-06-88	JP-T- 2500950	05-04-90
		WO-A- 8803559	19-05-88
EP-A- 0347078	20-12-89	AU-A- 3695589	05-01-90
		WO-A- 8912098	14-12-89
		GB-A- 2228007	15-08-90
		JP-T- 2504472	20-12-90